# Automated Multiple Point Stimulation Technique for Motor Unit Number Estimation

by

Marzieh Abdollahi

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2007

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

# Abstract

Motor unit number estimation (MUNE) is an electrodiagnostic procedure used to estimate the number of MUs in a muscle. In this thesis, a new MUNE technique, called Automated MPS, has been developed to overcome the shortcomings of two current techniques, namely MPS and MUESA. This method can be summarized as follows. First, a muscle is stimulated with a train of constant intensity current pulses. Depending on various factors, one to three MUs activate probabilistically after each pulse, and several responses are collected. These collected responses should be divided into up to $2^n$ clusters, such that each cluster represents one possible combination of $n$ Surface-detected Motor Unit Potentials (SMUPs). After clustering the collected responses, the average response of each cluster is calculated, the outliers are excluded, and similar groups are merged together. Then, depending on the number of response set groups, a decomposition technique is applied to the response clusters to obtain the $n$ constituent SMUPs. To estimate the number of MUs, the aforementioned process is repeated several times until enough SMUPs to calculate a reliable mean-SMUP are acquired. The number of MUs can then be determined by dividing the maximal compound muscle action potential (CMAP) size by the mean-SMUP size. The focus of this thesis was on using pattern recognition techniques to detect $n$ SMUPs from a collected set of waveforms.

Several experiments were performed using both simulated and real data to evaluate the ability of Automated MPS in finding the constituent SMUPs of a response set. Our experiments showed that performing Automated MPS needs less experience compared with MPS. Moreover, it can deal with more difficult situations and detect more accurate SMUPs compared with MUESA.

# Acknowledgements

This is an opportunity for me to express gratitude to those people who supported me in one way or another. The first person I would like to thank deeply is my supervisor, Prof. Daniel Stashuk, from whom I have learned a great deal. His breadth and depth of knowledge, hardworking spirit, and experience has always been an inspiration for me. His constant support and guidance led my research toward the right direction and helped me overcome the obstacles.

I would also like to thank Prof. Calamai and Prof. Clausi for spending their precious time on reviewing my thesis and for their comments.

My profound gratitude and love goes to my husband, Hossein. I would like to thank him for all the joy and happiness he brought into my life and for his patience, especially in the phases of my intense work on this thesis. I cannot imagine how I could finish this work without his support.

I feel a great sense of gratitude for my parents, my sister, Maryam, and my brother, Majid. Without their support and encouragement, I could not achieve success in my life and work.

# Contents

# List of Tables

# List of Figures

# List of Abbreviations

ALS amyotrophic lateral sclerosis

AP action potential

CMAP compound motor action potential

IP interface pattern

IS incremental stimulation

MFP muscle fiber potential

MN motor neuron

MPS multiple point stimulation

MU motor unit

MUESA motor unit number estimation based on stochastic activation

MUNE motor unit number estimation

MUP motor unit potential

MUPT motor unit potential train

NMJ neuro muscular junction

SMA spinal muscular atrophy

SMUP surface-detected motor unit potential

STA spike triggered averaging

# Chapter 1

# Introduction

In this thesis, the biosignals resulting from electrical stimulation of human muscles are analyzed. Skeletal muscles are stimulated with an electrical current. The muscles react to this stimulation and produce electrical signals. These signals are processed using pattern recognition techniques and physiological facts. Some useful information about the muscles (e.g., information about the health of a muscle) can be obtained from this process. This chapter presents an overview to the physiology of human muscles and the nervous system, as well as the muscle contraction mechanism.

## 1.1 The Structure of Muscles and Nerves

Each movement in our body, either voluntary or involuntary, is caused by the relaxation or contraction of muscles. Human muscles convert chemical energy into mechanical form to produce force and cause motion. Voluntary contraction of muscles, such as eye movement, causes motion in a part of our body. However, most of the muscle contractions in our body, such as the contraction of our heart, are involuntary. Human muscles can be divided into the following groups.

**Cardiac muscle** or heart muscle, which is found in the heart, contracts involuntarily in order to pump the blood through our circulatory system. This contraction is self-excitable and does not require an impulse from the nervous system. The contraction rate of a cardiac muscle is about 72 times per minute.

**Smooth muscle** which is found in the walls of hollow organs of the body except for the heart, and contracts involuntarily under the control of the autonomic nervous system. The contraction of a smooth muscle controls the flow of liquids in hollow organs. For example, it controls the movement of blood in arteries; expels urine from the urinary bladder, and regulates the flow of air through the lungs [1].

**Skeletal muscle** is usually connected to the skeleton at one end or both. When a skeletal muscle contracts, some force is applied to bones and joints causing a movement in a part of our skeleton. Skeletal muscles usually contract voluntarily, but they can also contract involuntarily through reflexes.

Approximately 40% of our muscles are skeletal. The work in this thesis is only related to skeletal muscles. Therefore, from now on, the term *muscle* is used instead of *skeletal muscle*.

A muscle fiber is a single muscle cell, and the structural unit of contraction. An individual muscle contains hundreds to thousands of muscle fibers bundled together inside a connective tissue covering. Each muscle fiber has many myofibrils; myofibrils are bundles of filaments that run through the entire length of the muscle and are connected to the cell surface membrane at both ends. The length and diameter of a muscle fiber can vary between a few millimeters to 30cm and from 10-100$\mu$m respectively. When a muscle contracts, the length of the muscle fibers can reduce to 57% of their resting length [4].

## 1.2 Muscle Fiber Innervation and Contraction

The movement of skeletal muscles is controlled by the nervous system via the sending and receiving of messages through nerve fibers. When a message is sent, positive sodium ions flow toward the inside of the axon membrane, which causes the depolarization of the membrane. Then, potassium ions flow toward the exterior of the cell, which leads to the depolarization of the membrane. This disturbance in the resting potential of a nerve fiber membrane produces an Action Potential (AP). The return of the stimulated points to their resting potentials causes the stimulation of adjacent points and their depolarization. This process is continued until the message is propagated to the end of the fiber.

The procedure of muscle contraction begins when APs, which originated at $\alpha$ motor neurons in the spinal cord, are received from the neuron axons. Each axon is divided into several axonal twigs near its end. Each axonal twig is connected to a muscle fiber in the middle of the fiber. This junction is called a Neuro-Muscular junction (NMJ). When the AP propagating along the axon reaches the NMJ, acetylcholine is released. This release leads to the generation of an AP on the muscle fiber membrane. The AP propagates along the muscle fiber and into myofibrils causing the release of calcium ions. The release of calcium ions makes the fiber filaments slide together. This is when the muscle contracts [29, 20].

## 1.3 Motor Unit

The axon of an $\alpha$ motor neuron (MN) is connected to, from several, to hundreds of muscle fibers stimulating them simultaneously. The set of connected muscle fibers, the $\alpha$ MN in the spinal cord, and its axon is called a Motor Unit (MU). The muscle fibers of a MU are not adjacent; they are distributed across the cross section of a muscle. In a healthy muscle, the distribution of fiber diameters in a MU is a Gaussian distribution with a mean of 55 $\mu$m and standard deviation of 9 $\mu$m. The number of MUs in a muscle and the number of fibers in a MU (innervation ratio) depend on the functionality of the muscle. A small muscle such as an eye muscle, which needs precise control of motion, has hundreds of small MUs, each of them innervating a few fibers, whereas a large muscle such as a thigh muscle has a smaller number of MUs, each of them having hundreds of fibers [20, 13].

As mentioned in Section 1.2, the initiation of APs and their propagation along the muscle fibers during a contraction causes the depolarization of the muscle fibers. This depolarization generates an electric field in the fluid surrounding the muscle fibers, which can be detected by inserting a needle electrode into the muscle or by a skin surface electrode. The detected signal generated by a single muscle fiber is called a muscle fiber potential (MFP). The combination of muscle fiber potentials from all muscle fibers of a single motor unit is called a motor unit potential (MUP). In order to maintain force or increase the force, all muscle fibers of a MU are fired repeatedly, generating a train of MUPs called a motor unit potential train (MUPT). The number of MUPs in a MUPT depends on the firing frequency of the corresponding MU [18, 36].

### 1.3.1 Motor Unit Recruitment

Motor unit potentials are referred to as all-or-nothing signals. When a muscle contracts, each MU will or will not be activated. If the level of contraction is higher than a threshold (recruitment threshold), the MU is activated. The activated MU remains active as long as the level of contraction remains above the recruitment threshold. When a MU becomes active, all of its muscle fibers are contracting with maximum force. Increasing the level of contraction in this situation does not increase the detected potential of that individual MU. However, increasing the contraction level increases the firing rates and number of activated MUs. According to the Henneman size principle [26], as the contraction level increases, the MUs are recruited in the order of their sizes. Smaller MUs are recruited even with a small contraction, while larger MUs need a large contraction to be activated [2].

### 1.3.2 Motor Unit Number Estimation

Contractions of different muscles in our body are the results of the activation of different combinations of MUs. Therefore, MUs, which are the functional units of muscles, have an important role in muscle physiology. The number of MUs in a muscle can be affected by aging or several diseases. A good estimation of the number of MUs can help in the diagnosis of neuromuscular disorders, evaluating the progress of disorders that result in the progressive loss of MUs, as well as assessing the responsiveness of patients suffering from the above mentioned disorders to different treatments.

Motor unit number estimation (MUNE) is an electrodiagnostic procedure used to estimate the number of motor units (MUs) in a muscle [8]. In this research, the focus is on different MUNE techniques and introducing an improvement to an existing technique. The next section explains how useful waveforms can be recorded from a muscle.

## 1.4 EMG

An electromyographic (EMG) signal, which is the whole electrical potential detected from a muscle during a contraction, is formed by the combination of MUPTs generated by several MUs. Electromyography is the medical procedure in which an EMG signal is recorded.

EMG signal characteristics depend on several factors including the shape and size of the electrodes used for EMG recording, the position and orientation of the electrodes relative to active MFs, the strength of the contraction, the impedance and temperature of the tissue, the control scheme of the peripheral nervous system, and the anatomical and physiological features of the active MUs. Clinical electromyography can provide useful information about the relationship between EMG signals and the physiological and morphological features of the involved MUs and muscle fibers [11, 13]. EMG studies are divided into two categories. The first group studies EMG signals of healthy muscles, while the second group analyzes EMG signals of diseased, aged or fatigued muscles. Neuromuscular disorders change the structure of muscles and affect how they function resulting in abnormal EMG signals. Analyzing EMG signals of diseased muscles is a diagnostic tool for characterizing neuromuscular disorders [36, 35].

Using different types of electrodes (micro or macro needle electrodes or surface electrodes) produces different types of EMG signals. An EMG signal detected from a small surface primarily represents the contributions of close MFs, while an EMG signal detected from a large surface shows the contributions of both close and far MFs [29, 35]. In this research, surface electrodes are used for collecting EMG signals from muscles.

There are two major approaches for detecting the EMG signals of MNs. The first approach examines signals collected from voluntarily contracted muscles. When the contraction is weak, individual MUPs can be detected easily either manually or automatically. However, these MUPs represent smaller MUs and cannot represent all MUs of a muscle. On the other hand, stronger contractions lead to the firing of larger MUs, as well as increasing the firing rates of previously recruited MUs. In this case, the MUPs of different MUs overlap and run into each other resulting in a very complex signal called a composite EMG signal or Interface Pattern (IP). The detected EMG signal represents the activity of many MUs which are firing asynchronously. IP analysis can provide some useful information about the number of MUs, as well as the firing rates and characteristics of recruited MUs. However, IP analysis does not provide direct information about individual MUPs.

It only allows the analysis of the effect of the superpositions of MUPs.

In the second approach, which is used in this research, there is no voluntary contraction; the muscle is subjected to an external electrical stimulation. This external stimulation synchronizes the activation of MUs producing bi-phasic or tri-phasic potentials. In this approach, it is not possible to control the size of the MUPs. However, the number of activated MUs can be controlled by changing the stimulus intensity. Although obtaining individual MUPs and the number of MUs in this approach seems easy, it has some practical and theoretical limitations [36, 34]. These limitations will be discussed further in Chapter 2.

## 1.5 Neuromuscular Disorders

Neuromuscular disorders are muscle-related and nerve-related disorders that disturb the normal structure and functionality of the muscles. Neuromuscular disorders can be divided into the following categories.

1. **Myopathies** in which the muscle fibers do not function normally, resulting in muscular weakness. Other symptoms of myopathy include muscle aching, cramping, pain, stiffness, tenderness and tightness. Myopathies lead to several abnormalities in the structures of muscle fibers, such as unusual decreases in the sizes of muscle fibers, replacement of muscle fibers with fatty tissue, splitting of muscle fibers into several thinner muscle fibers, etc,.

2. **Neuropathies** in which the motor neurons, which are a part of body's nervous system, are affected. The early sign of neuropathic disorders is the loss of MNs. When a MN dies, the connections of its muscle fibers with the nervous system are lost,

and the muscle fibers are denervated. Then, the adjacent unaffected MNs generate collateral nerve sprouts and re-innervate the orphaned muscle fibers [18].

As mentioned in Section 1.4, since neuromuscular diseases cause some abnormalities in EMG signals, quantitative EMG analysis is a strong diagnostic tool for characterizing neuromuscular disorders. Moreover, analyzing the size, shape and firing pattern of MUPs collected during electrodiagnostic testing can help physicians distinguish a myopathy from a neuropathy. There are other diagnostic methods for detecting neuromuscular disorders including physical examination, muscle biopsy, and laboratory examinations (protein synthesis, genetic testing) [18, 6]. Clinical EMG however, allows the function and physical layout of the muscle fibers of a muscle to be studied.

## 1.6 Thesis Organization

The objective of this research is to improve existing techniques for estimating the number of MUs in a muscle. In this chapter, an overview to the physiology of human muscles and the nervous system was presented. Then, the contraction mechanism in which the chemical energy is converted to mechanical energy was described, and electromyographic procedures were explained. The rest of the thesis is organized as follows.

In Chapter 2, the importance of MUNE study will be explained. Then, a brief review of existing techniques for estimating the number of MUs will be presented to give a better understanding of the project.

The first and second steps of our proposed MUNE technique, called Automated MPS, are described in Chapter 3 and 4 respectively. In the first step of Automated MPS, the muscle is stimulated with a train of constant intensity current pulses, and several responses are collected while $n$ MUs are intermittently responding. The collected signals are divided into up to $2^n$ clusters, so that each cluster represents one possible combination of $n$ S-MUPs. Then, the cluster representative of each cluster is calculated. In the second step of Automated MPS, the $n$ constituent Surface-detected Motor Unit Potential (SMUPs) are extracted. These SMUPs can then be used to calculate the mean-SMUP and estimate the number of MUs. Chapter 4 explains the decomposition process either when all combinations of the $n$ SMUPs are observed, or when one of the combinations is absent.

Chapter 5 presents the experiments performed to evaluate the ability of Automated MPS in finding SMUPs. Corresponding to the results of the experiments, a brief discussion is presented in Chapter 6. Finally, Chapter 7 highlights important contributions of this research and suggests some directions for future work in this area.

# Chapter 2

# Background and Related Work

In this chapter, a short but comprehensive survey of the existing MUNE techniques is presented. Section 2.1 explains why motor unit number estimation is important. Section 2.2 describes the required information for estimating the number of MUs in the muscle and divides the MUNE techniques into two categories based on the contraction type they employ (voluntary contraction versus external stimulation). Section 2.3 reviews the MUNE techniques in which there is no voluntary contraction; the SMUPs are collected while the nerve is stimulated with an external electrical stimulation. Section 2.4 studies the MUNE techniques in which the SMUPs are collected while the muscle is contracting voluntarily. Section 2.5 compares different MUNE techniques based on their ability to find the mean-SMUP.

## 2.1 The Importance of Motor Unit Number Estimation

There are several motivations for developing an accurate and reliable method for estimating the number of MUs in a muscle. MU studies can provide useful information about the structure of the human brainstem and spinal cord, as well as the intervention of muscles. As mentioned in Section 1.5, several factors such as aging and some neuromuscular disease can affect the MUs. Due to the atrophy of individual muscle fibers in elderly people, their muscle bulk is usually less than young people. MUNE techniques are good tools for studying the effect of aging on the population of MUs. On the other hand, the number of MUs decreases progressively in muscle denervating disorders such as Amyotrophic Lateral Sclerosis (ALS), Spinal Muscular Atrophy (SMA), poliomyelitis and several inherited and acquired peripheral neuropathies. A good estimation of the number of MUs allows changes in the number of MUs in the muscle to be detected, helping to diagnose the above mentioned disorders during early stages of disease involvement. MUNE also allows the severity and history of muscle denervations in such disorders to be monitored, and the evaluation of the responsiveness of patients to different therapies [8].

During the past 40 years, different MUNE techniques have been developed for the above mentioned purposes. The next section explains the information required for estimating the number of MUs in a muscle.

## 2.2 Electrophysiological Motor Unit Number Estimation Techniques

The first electrophysiological MUNE technique, called the Incremental Stimulation technique (IS), was developed in 1971 [28]. After that, researchers have proposed several other MUNE techniques. Although each method is different from other methods, all of them are based on the same underlying premise. In all MUNE techniques, the following two pieces of information are needed:

1. The summated response of all of the MUs in the muscle, called a Compound Motor Action Potential (CMAP); to get a CMAP, the nerve is stimulated at a super-maximal level resulting in activating all MUs at approximately the same time. The resulting evoked potential, which is recorded with a surface electrode, is the CMAP. If the stimulation is repeated several times, a very similar biphasic signal can be obtained.

2. The average surface-detected response of all MUs, called the mean-SMUP; the usual way to get a mean-SMUP is to study a small number of SMUPs, provided that these SMUPs can represent the responses of the entire MU population. Different MUNE techniques calculate the CMAP in the same way, but they estimate the mean-SMUP in different ways. The existing methods for obtaining the mean-SMUP are divided to two categories based on the type of muscle contraction they employ. One group of MUNE techniques are based on the external stimulation of muscle, while the other group is based on the voluntary contraction of muscle.

Having the above signals (CMAP and mean-SMUP), the number of MUs of a muscle is calculated according to Equation 2.1 [34, 15].

$$\text{Number of MUs} = \frac{\text{(amplitude or area) of the maximum CMAP}}{\text{(amplitude or area) of the mean-SMUP}} \tag{2.1}$$

The above estimation of MU numbers is valid only if the same electrode configuration is used for detecting the maximal CMAP and mean-SMUP. Moreover, the SMUPs used for estimating the mean-SMUP should represent the activity of different MUs of the muscle [33].

## 2.3 External Stimulation-based MUNE Techniques

The MUNE techniques described in this section estimate the mean-SMUP while the nerve is stimulated with an external electrical stimulation.

### 2.3.1 Incremental Stimulation (IS)

In 1971, McComas et al. [28] described the first electrophysiologic MUNE technique, called Incremental Stimulation (IS). The IS technique was originally applied to the extensor digitorum (EDB) muscle and then to intrinsic hand muscles [9]. This technique uses the fact that different MUs have different activation thresholds to stimulus intensity. Therefore, if the stimulus intensity is increased gradually, each observed increment in the CMAP can be attributed to the activation of an additional single MU. After about 10 increments, the mean-SMUP is calculated by dividing the amplitude of the last observed CMAP, which is the largest CMAP, by the number of increments. Then, the nerve is stimulated at the maximal level to get the maximum CMAP, generated by all of the MUs of the muscle, and the number of MUs is calculated using Equation 2.1. The validity of the IS technique depends on the following assumptions.

1. The observed increase in the CMAP at each step is the result of the activation of a single new MU.

2. The SMUPs used for calculating the mean-SMUP represent the responses of all MUs.

3. The recorded signals are coming only from the muscle being tested.

**Alternation Phenomenon:**

Applying the IS technique to human muscles is possible [16, 23, 5], but not easy. The main issue of the IS technique is the alternation phenomenon. The firing threshold of each MU is not a fixed number. There is a range of stimulus intensity over which the probability of MU recruitment increases from 0 to 100%. Considering the fact that the number of MUs in a muscle can be a large number, it is very likely that the thresholds of several MUs overlap with each other even if the stimulus intensity is not very large. When the thresholds of $n$ MUs overlap at a specific stimulus intensity, $2^n - 1$ incremental steps might be observed if a sufficient number of stimuli are applied. Each of these steps shows one combination of $n$ SMUPs. The above phenomenon is called alternation. The number of alternations can be larger than the number of MUs involved. For example, 14 combinations can result from the alternation of only four MUs. Alternation can lead to the underestimation of the mean-SMUP size and overestimation of the number of MUs [17, 14].

To solve this problem, some modifications were made to the original IS techniques [3, 19, 31]. Moreover, other stimulation-based techniques such as multiple point stimulation (MPS) [17] and F-respone [37], as well as some voluntary contraction-based techniques such as spike-triggered averaging (STA) [10, 39, 7] were developed in order to collect samples of SMUPs, free of alternation.

### 2.3.2 Multiple Point Stimulation (MPS)

MPS [17] is a stimulation-based MUNE technique that can collect several SMUPs without alternation. In this technique, the motor nerve is stimulated at several sites, and one SMUP is collected at each site. This SMUP is the response of the first activated MU at that site. The aforementioned process is repeated several times until enough SMUPs (about ten) are collected for finding a meaningful mean-SMUP. Since 50-100 mm of the motor nerve should be accessible in MPS, distal muscles such as the median innervated thenar group, the ulnar innervated hypothenar group, the first dorsal interosseous/adductor pollicis muscle group, and the extensor digitorum brevis muscle are the best choices. The MPS technique is based on several assumptions. The first two are the same as those of IS. At each site, only one MU should be activated, and the collected SMUPs should be representative of the responses of the entire MU population. Furthermore, it is assumed that enough SMUPs (according to [16], about 10 SMUPs) can be collected for calculating the mean-SMUP.

The most important advantage of MPS is its ability to calculate an unbiased mean-SMUP that can represent the responses of all MUs. This claim is based on several observations. First, the collected SMUPs have a wide range of sizes and relative latencies. Second, the size distribution of SMUPs of individual experiments can reflect the distribution of pooled subjects. Third, similar thenar MU twitch tensions can be obtained with MPS and other techniques such as STA. Fourth, the mean-SMUP and MUNE results of MPS and other MUNE techniques are similar. Finally, the possibility of exciting the same axon at two sites is slightly higher than the chance of exciting two different MUs [15, 16]. Other advantages and disadvantages of MPS compared with other MUNE techniques are as follows.

**Advantages of MPS:**

1. The mean-SMUP in MPS is not calculated based on a statistical estimation or using a modified algorithm for alternation correction, but rather based on the averaging of real SMUPs.

2. There is no alternation in MPS.

3. Stimulation near the motor thresholds are well tolerated by subjects.

4. Besides collecting the SMUPs, some other pathological properties of MFs can be obtained using MPS. For example, MPS can detect the reduction of the SMUP size in response to repetitive stimulations. This size reduction happens in ALS and other rapidly developing axonal neuropathies. The inability of methods such as IS in detecting this reduction can cause some errors in the estimation of the number of MUs.

**Disadvantages of MPS:**

1. Obtaining all SMUPs from one site is not possible.

2. MPS cannot be applied to proximal muscles.

3. The most important disadvantage of MPS is that the operator needs to have considerable skill and experience to be able to collect a sufficient number of SMUPs in a reasonable time (about 20 min). The MPS operator should have the following skills:

   - Recognizing the individual SMUPs as all-or-nothing signals, as well as recognizing alternation or other errors that prevent from identifying the stimulation of a single MU;

- Having indispensable sensitivity for performing MPS (e.g., the operator should change the position and the pressure on stimulating electrodes accordingly when a change happens in the detected surface-EMG signal in response to a small change in the stimulus intensity);
- Learning how to search along the course of the nerve in order to stimulate low threshold MUs; the operator should make small changes in the stimulus intensity and move the electrodes distally or proximally along the course of nerve. When a suitable SMUP is found, the operator should learn to hold the electrode steadily enough to get a good SMUP.

An experienced operator can acquire higher threshold SMUPs at the same site or acquire SMUPs while alternation happens. Finding the individual SMUPs is more difficult in young and healthy people compared with old people or patients with various neuropathies. The reason is that the number of motor nerve fibers in young/healthy people is large, and their thresholds overlap a lot, while the number of motor nerve fibers in old/diseased people is much smaller, and their thresholds are well separated [15].

### Adapted MPS

Adapted MPS [40, 42] is a MUNE technique based on the IS and MPS techniques. In this method, incremental stimulations are applied to different sites along the median nerve, and two or three SMUPs are collected at each site. However, the SMUPs are used for calculating the mean-SMUP only if they are alternation-free. Adapted MPS is used when the IS method is not applicable because of alternation, or when finding 10 SMUPs is difficult with MPS. It is a fast non-invasive technique, which does not need a specific recording system. However, this technique requires an experienced operator, and allows the possibility of recognizing the same SMUP as a new one [41, 42].

### 2.3.3 Statistical MUNE

In most of the MUNE techniques, several single SMUPs are collected using different techniques and are averaged to get the mean-SMUP. Another approach for finding the mean-SMUP is to estimate the mean-SMUP size indirectly using the statistical characteristics of a sequence of CMAPs. The MUNE statistical technique uses Poisson statistical assumptions, in which it is assumed that all samples have the same size, the size of each sample does not change, and the histogram of all samples is skewed to the right.

As mentioned previously, the activation threshold of each MU is not a finite value. The activation threshold of a single MU is a range in which the firing probability of the MU increases as the stimulus intensity increases. This range can be shown as a sigmoid curve. After each stimulus, each of the MUs of the muscle will fire with a probability depending on the stimulus intensity. If the stimulus intensity is very low, only MUs with the lowest thresholds are activated. In this case, after some of the stimuli, none of the MUs will fire resulting in a zero CMAP (just the baseline). In this case, the variance of the responses can be calculated by dividing the SMUP size by the Poisson assumptions. If the stimulus intensity is higher, different combinations of several SMUPs are superimposed on the baseline. For $n$ recorded sequential responses, the variance of the responses and the estimated mean-SMUP size can be calculated as follows.

$$C = [C_1, ..., C_n] \tag{2.2}$$

$$\sigma_C^2 = \frac{\sum_{i=1}^{n} C_i^2 - \frac{\sum_{i=1}^{n} C_i^2}{n}}{n-1} \tag{2.3}$$

$$S = \frac{\sigma_C^2}{\mu_C - \min_C} \tag{2.4}$$

where $C$ is the submaximal CMAP and $S$ is the mean-SMUP size. To derive the MUNE, the maximal CMAP size should be divided by $S$. A MUNE based on statistical analysis has the following advantages and disadvantages.

**Advantages of MUNE based on statistical analysis:**

1. Statistical MUNE does not need a skilful operator.

2. This method is fast, efficient and reproducible.

3. Unlike most MUNE techniques, alternation is not an issue in statistical MUNE.

**Disadvantages of MUNE based on statistical analysis:**

1. Statistical MUNE needs patient cooperation. If the patient does not cooperate, deriving the MUNE will be time consuming.

2. This method cannot be applied to proximal muscles.

3. Several variables such as stimulation current, recording window levels, recording window size, and recording window number should be specified by the operator [11, 12].

**Bayesian Statistical MUNE**

The Bayesian statistical MUNE technique has been developed to improve the deficiencies of the statistical MUNE. In the statistical MUNE, it is assumed that the numbers of alternating MUs have a Poisson distribution, and each SMUP has a fixed and identical size. These assumptions are not always true. The Bayesian approach incorporates threshold variability, the variability between/within SMUPs, and baseline variability. This method uses Markov chain Monte Carlo methods to provide useful information about the population of MUs, as well as individual MUs. Up to now, this technique has been used for deriving the MUNE of diseased muscles in which the number of MUs are small; it cannot be applied to healthy muscles with a large number of MUs [33].

### 2.3.4 F-Response

The F-response technique is another stimulation-based MUNE technique that does not suffer from the alternation problem [37]. Similar to MPS, the F-response technique is best applied to distal muscles. As mentioned in Section 2.3.2, about 10 SMUPs are required for calculating a meaningful mean-SMUP. To obtain 10 distinct SMUPs using F-responses, at least 200-300 stimuli should be applied to the motor nerve. Depending on the probability of an F-response, trains of stimuli should be applied to one, two or sometimes three sites to collect enough SMUPs for deriving the mean-SMUP. Another strategy is to estimate the number of MUs two or three times and average them [19].

After stimulating the nerve, the collected F-responses are grouped based on their sizes, shapes and latencies. The F-responses with identical size, shape, and latency that are observed twice or more result from the firing of one MU, while F-responses that are observed only once are considered either as compound F-responses (i.e., they represent the combination of more than one SMUP) or F-responses resulting from the firing of MUs with very low probability of activation. The grouping of F-responses can be either manual (finding the identical F-responses visually) or automatically (using template matching algorithms). In both cases, after collecting about 10 SMUPs, the mean-SMUP is calculated by point by point averaging, and the number of MUs is calculated using equation 2.1.

The validity of MUNE using samples of F-responses depends on the following assumptions.

1. A sample of SMUPs that are the responses of individual MUs can be obtained from the F-responses.

2. The sample can represent the relative number of SMUPs with different shapes, sizes and latencies.

3. The collected SMUPs are independent from the stimulus intensity used for collecting the F-responses.

In a healthy muscle, the probability of observing an F-response for each motor unit is much less than 10% (about 1 in 50 or 1 in 100 for the median innervated thenar muscle). Since the probability of observing the F-response of one MU is independent of that of another MU, the probability of observing a combination of two MUs in the F-response is the product of the probability of observing each of them, which is a very small number. For example, if the probability of observing each of the two F-responses is 5% (10-15 times in 200-300 stimuli), the probability of observing the combination of them is 0.25% ( once in 300 stimuli). Therefore, in most cases, it can be assumed that if an F-response is observed more than once in 200-300 stimuli, it is coming from only one MU. The only exception is in some pathological conditions such as ALS in which the F-response of each MU has an observation probability larger than 50%. In these cases, assuming that F-responses that are observed more than once resulted from the activation of only one MU can cause errors in the mean-SMUP and the estimated number of MUs.

The main advantage of the F-response method is that it does not need a skillful operator to perform the experiments. It also needs minimal patient cooperation. Another advantage of this method is that it can provide additional information about the latency and conduction velocity of axons. The main disadvantage of the F-response is that it is a time consuming method, which can be inconvenient for subjects (because 200-300 stimuli should be applied to the nerve). Another disadvantage is that it cannot be applied to proximal muscles and needs special software; it cannot be performed using usual surface-EMG devices [15].

### 2.3.5 MUNE Based on Stochastic Activation (MUESA)

Most of the MUNE techniques avoid alternation, because they want to collect SMUPs directly when one single MU is activated. However, several MUNE techniques, reported in the literature, use the alternation phenomenon. In these techniques, the signals are collected from the muscle while several MUs are alternating. The statistical method [11, 12] and MUNE based on stochastic activation (MUESA) [34] are examples of such techniques. In the MUESA method, the nerve is stimulated with a train of constant-intensity and several responses (about 40) are collected while $n$ MUs (1-3 MUs) are alternating. The collected signals are grouped using a hierarchical clustering technique such that similar waveforms are in the same group. Each group shows one combination of the $n$ SMUPs. For $n$ activated MUs, up to $2^n$ combinations might be observed. If a full set of combination ($2^n$ groups) are observed, the constituent SMUPs are extracted based on their sizes or relative firing rates. MUESA assumes that the size of the combination of two SMUPs is larger than the size of each of them. Therefore, the largest observed signal shows the combination of the $n$ SMUPs and the smallest signal shows the case in which none of the $n$ MUs are firing. For two alternating MUs, MUESA finds the two single SMUPs by subtracting the second and third largest signals from the largest one. Another way for extracting the two SMUPs is to subtract the most prevalent signal (the one that has been observed the most) from the second and third most prevalent signals. On the other hand, if the set of responses is not complete, and some combinations of SMUPs are not observed, the constituent SMUPs can be extracted based on their relative firing rates (i.e., how may times each response has been observed) provided that some groups are obviously prevalent.

After extracting one, two or three SMUPs from one site, the same process is performed several times at other sites along the muscle until enough SMUPs (about 10) are collected for calculating the mean-SMUP and deriving the MUNE using equation 2.1 [34].

The following paragraphs review the advantages and disadvantages of the MUESA method.

**Advantages of MUESA:**

1. The main advantage of MUESA method is that unlike most of the MUNE techniques, MUESA does not need to assume that each observed signal represents the activation of a single MU. MUESA can be done in the presence of alternation, which is one of the most serious problems of other MUNE techniques.

2. Applying constant-intensity stimulations to the nerve leads to the probabilistic activation of a limited number of MUs after each pulse. Moreover, the stationarity of the muscle can be examined using constant-intensity stimulations. Most MUNE techniques assume that the response of each MU does not change during the procedure. However, this assumption is not true in some diseases such as ALS. In such a disease, if the same stimulation is repeated several times, the amplitude of the CMAP and its constituent SMUPs decreases. The detection of these non-stationarities is important; MUESA is able to do that.

3. The calculated mean-SMUP in MUNE is not estimated; it is the real average of SMUPs [34].

**Disadvantages of MUESA:**

1. Although the MUESA operator does need to be as skilful as an MPS operator, he/she should be experienced enough to be able to set the stimulation such that one, two or three MUs alternate [34].

2. MUESA assumes that the size of the combination of two SMUPs is always larger than the size of each of them. This assumption is not always valid. The combination of two SMUPs can be smaller than each or both of the SMUPs if there are some phase cancellations as the waveforms summate.

3. The other approach of MUESA for extracting the constituent SMUPs of a combination of several SMUPs is based on the estimation of the firing rates of the MUs. The estimation of the probability of observing each combination of the $n$ activated SMUPs can be very different from the number of times it is actually observed in the experiments. Hence, subtracting the second most prevalent signal from the most prevalent signal does not necessarily give one of the SMUPs even if two groups are obviously more prevalent.

4. MUESA cannot be applied to proximal muscles.

## 2.4 Voluntary MUNE Techniques

The voluntary MUNE techniques are similar to stimulation-based techniques in that the number of MUs is estimated by dividing the CMAP size by the mean-SMUP size. However, in contrast to the stimulation-based techniques, in voluntary techniques, the MUs are activated by voluntary muscle contractions, and single MUs are detected using an intramuscular needle electrode. The most well-known voluntary MUNE technique, Spike

Triggered Averaging (STA) [7, 10, 39], is explained in the next section.

### 2.4.1 Spike Triggered Averaging (STA)

In this technique, weak voluntary contractions are used to activate a MU, and the signals are recorded simultaneously from the surface electrode and intramuscular needle electrode. According to volume conduction theory, the potential of each MU will be detected by both electrodes at the same time. The needle electrode detects the activation of a MU by detecting its MUP. Each MUP is isolated by adjusting a window or level threshold. If a window threshold is used, the MUP is detected when its amplitude is between a low and high limit. If a level threshold is used, the MUP is detected when its amplitude exceeds an adjustable threshold. The times of MUP detections are used as triggers for selecting time-aligned sections of the surface EMG signal. About 100-200 surface EMG signals are averaged then to estimate the SMUP, which represents the activation of the MU. Then, the needle electrode is moved to another site to estimate another SMUP. This procedure is repeated until 15-20 SMUPs are collected for finding the mean-SMUP. Since only one MUP can be obtained from one site, the STA should be applied to 15-20 sites in order to get enough MUPs [7].

Decomposition-enhanced STA is a variant of STA that can collect 3-10 SMUPs from one site during a single contraction. In decomposition-enhanced STA, a series of signal processing and pattern recognition techniques are used to decompose the EMG signal detected by the needle electrode to determine the activation times of MUs. The MUPs of a specific MU in the EMG signal are detected and classified to be used as triggers for averaging the selected sections of the surface EMG signal and estimating the SMUP of that MU [38].

The following paragraphs compare STA and Decomposition-enhanced STA, as well as with other MUNE techniques.

**Advantages:**

- *STA:*

1. Since the contractions are voluntary, the MUs are not activated simultaneously. Therefore, STA does not suffer from the alternation problem.

2. STA can be applied to distal muscles, as well as proximal muscles.

3. The contraction can continue until obtaining a SMUP with a sufficient signal to noise ratio.

4. The effect of the synchronous activation of other MUs is small.

- *Decomposition-enhanced STA:*

1. Since the contractions are voluntary, the MUs are not activated simultaneously. Therefore, this method does not suffer from the alternation problem.

2. Decomposition-enhanced STA can be applied to distal muscles, as well as proximal muscles.

3. The operator needs less skill compared with the STA technique.

4. Several SMUPs can be obtained from one contraction. Hence, fewer contraction are required compared to the STA technique.

5. The consistency and accuracy of the trigger MUPs can be reviewed.

6. The contraction level can be higher than that of the STA technique. Therfore, larger MUs with higher recruitment thresholds can be studied.

7. Using the EMG decomposition results, other neurophysiological information can be obtained.

8. It is possible to confirm that each SMUP collected during one contraction is the response of a unique MU.

**Disadvantages:**

- *STA:*

1. STA needs needle electrodes.

2. STA needs patient cooperation. The patient should maintain a steady contraction during the signal collection.

3. The required time for MUNE is large, because only one SMUP is obtained from one site and one contraction.

4. Reviewing the accuracy and consistency of MUPs used as triggers is not possible.

5. The contraction level should be low to make sure that only one MU is activated. This leads to a possible sampling bias toward MUs with low recruitment thresholds.

6. It is sometimes not easy to establish the baseline and onset of the SMUPs. Hence, it is difficult to measure the SMUP area accurately.

7. It is possible that the same SMUP is sampled during different contractions.

- *Decomposition-enhanced STA:*

1. The required EMG signal decomposition algorithms are not widely available.

2. It is not easy to extract the activity of one MU from an EMG signal (which is a composite of the activities of several MUs) with a good signal to noise ratio.

3. If the collected SMUPs are noisy, defining the landmarks (onset, positive and negative peaks, and end) will be difficult. Operator interpretation can lead to large variability.

4. To get SMUPs with better signal to noise ratios, either very low contractions or contractions maintained for a long period of time (30-60 S) are required; this can cause some fatigue.

5. Compared with the STA technique, this technique can be biased towards larger SMUPs (MUs with higher recruitment thresholds) and lower MUNE values. Compared with other MUNE techniques, it can be biased towards sampling small MUs which leads to a larger MUNE value.

6. Since the SMUPs are collected during different levels of contraction, there might be different degrees of bias depending on the sampled MUs recruitment thresholds [38].

## 2.5 Conclusions

The ability to estimate the number of MUs in a muscle can provide useful information about the severity of muscle denervating disorders, as well as the course and responsiveness of these disorders to different treatments. It also can be used to study the effect of aging on the population of MNs. An ideal MUNE technique should be easily applicable to any muscle and in a reasonable amount of time. In addition, it should be sensitive to changes in the size or number of MUs in order to assess both myopathies and neuropathies. None of the existing MUNE techniques can achieve these goals. In all MUNE techniques, the number of MUs is estimated by dividing the maximal CMAP size to the mean-SMUP size provided that the mean-SMUP represents the entire population of MUs. The maximal CMAP is always obtained by stimulating the muscle such that all MUs are activated at the same time, and recording the resulting potential with a surface electrode. The mean-SMUP can be calculated or estimated a number of different ways.

The more established MUNE techniques include Incremental Stimulation, Multiple Point Stimulation, Statistical MUNE, Spike Triggered Averaging, and F-response. Each technique has several advantages and disadvantages depending on how, when and where it should be used. The most important criterion for comparing MUNE techniques is their ability to find an unbiased mean-SMUP that is representative of the responses of all MUs in the muscle.

The first MUNE technique, Incremental Stimulation, cannot derive a good MUNE in the presence of alternation. The alternation happens when the thresholds of several MUs overlap, resulting in a variable number of fired MUs in response to a stimulus.

The statistical technique assumes that the size of an SMUP does not change in response to repetitive stimulation. In some neurogenic disorders, the size of SMUPs decreases during repetitive stimulations causing unwanted variation in the CMAP and a poor estimation of the mean-SMUP.

Spike Triggered Averaging is performed during low voluntary contractions. Therefore, MU sampling is biased toward low threshold MUs (smaller MUs), which lead to underestimating the mean-SMUP and overestimating the number of MUs.

The F-response method assumes that if the same F-response occurs more than once with the same size, shape and latency, it comes from one MU. In some diseases, the chance of observing an F-response is abnormally high. This increases the risk of observing the combination of several SMUPs more than once resulting to a biased mean-SMUP.

The Multiple Point Stimulation technique is one of the best existing MUNE techniques, because it can find the least unbiased mean-SMUP for the MUNE. The reason for this is that in MPS, a wide range of SMUPs with different sizes and relative latencies are collected for calculating the mean-SMUP. Since collecting 10 unique SMUPs from a muscle in a reasonable time needs a very skilful operator, it is useful to develop an MPS-based MUNE technique that can be performed more easily by experts, as well as inexperienced operators. Automated MPS, proposed in this thesis, tries to achieve this goal. Automated MPS, is also similar to MUESA in some steps. In both MUESA and Automated MPS, several responses are collected from each site in which, one, two or three MUs alternate. The responses are sorted and clustered; then the responses are extracted to their constituent SMUPs. The aforementioned steps are repeated several times at several sites until a suitable number of SMUPs to calculate the mean-SMUP (about ten [16]) are collected. The difference between Automated MPS and MUESA is the clustering algorithm and the decomposition procedure they employ. The next two chapters explain the details of Automated MPS.

# Chapter 3

# Clustering the Response Set

As mentioned in Chapter 2, in all MUNE techniques, the number of MUs of a muscle is obtained by dividing the amplitude or area of the maximal CMAP by the amplitude or area of the mean-SMUP. Obtaining the maximal CMAP is straightforward; the nerve is stimulated at maximum level, and the detected surface potential is stored. The difficult part is finding the mean-SMUP. In Chapter 2, existing MUNE techniques were studied and compared with each other based on their ability in obtaining an unbiased mean-SMUP. It was concluded from this comparison that MPS can achieve this goal better than other MUNE techniques. However, MPS is not an ideal MUNE method, and it has several deficiencies. The most important disadvantage of this technique is that the operator needs to be experienced and skilful to collect enough SMUPs in a reasonable amount of time (about 20 minutes), and recognize alternation or other errors. The Automated MPS technique, introduced in this thesis, is a MUNE technique that can solve this problem; it can be performed by experienced operators, as well as operators that do not have a large amount of experience.

## 3.1 Basic Description of Automated MPS

In response to an external stimulation of a motor nerve, the MUs of the muscle discharge in a probabilistic manner. The major source of difficulty in finding a mean-SMUP in most MUNE techniques is this probabilistic activation (alternation). However, some techniques such as MUESA or Automated MPS work based on this phenomenon.

In the first step of Automated MPS, described in this chapter, the muscle is stimulated with a train of constant intensity current pulses, such that several responses are collected. Depending on various factors, one, two or three motor units are activated after each pulse. Consequently, for $n$ motor units, $2^n$ possible combinations of SMUPs might be observed in the dataset (in this work, the value of $n$ can be any number between one and five). In the first step, the responses should be divided into up to $2^n$ clusters, such that each cluster represents one possible combination of $n$ SMUPs. Then, the cluster representative of each cluster (the average response) is calculated and the second step of the algorithm (described in Chapter 4) in which $n$ constituent SMUPs are extracted, is performed.

The above-mentioned process is repeated several times until enough SMUPs (according to [**?**, **?**, 22] about 10) for estimating the mean-SMUP are collected. Then, the maximal CMAP is acquired, and the number of MUs is calculated by dividing the maximal CMAP amplitude or area by the mean-SMUP amplitude or area. Figure 3.1 describes different steps of Automated MPS.

## 3.2 Response Set

A MU can respond to a train of constant-intensity stimulation in the following ways.

- The MU may always be activated.

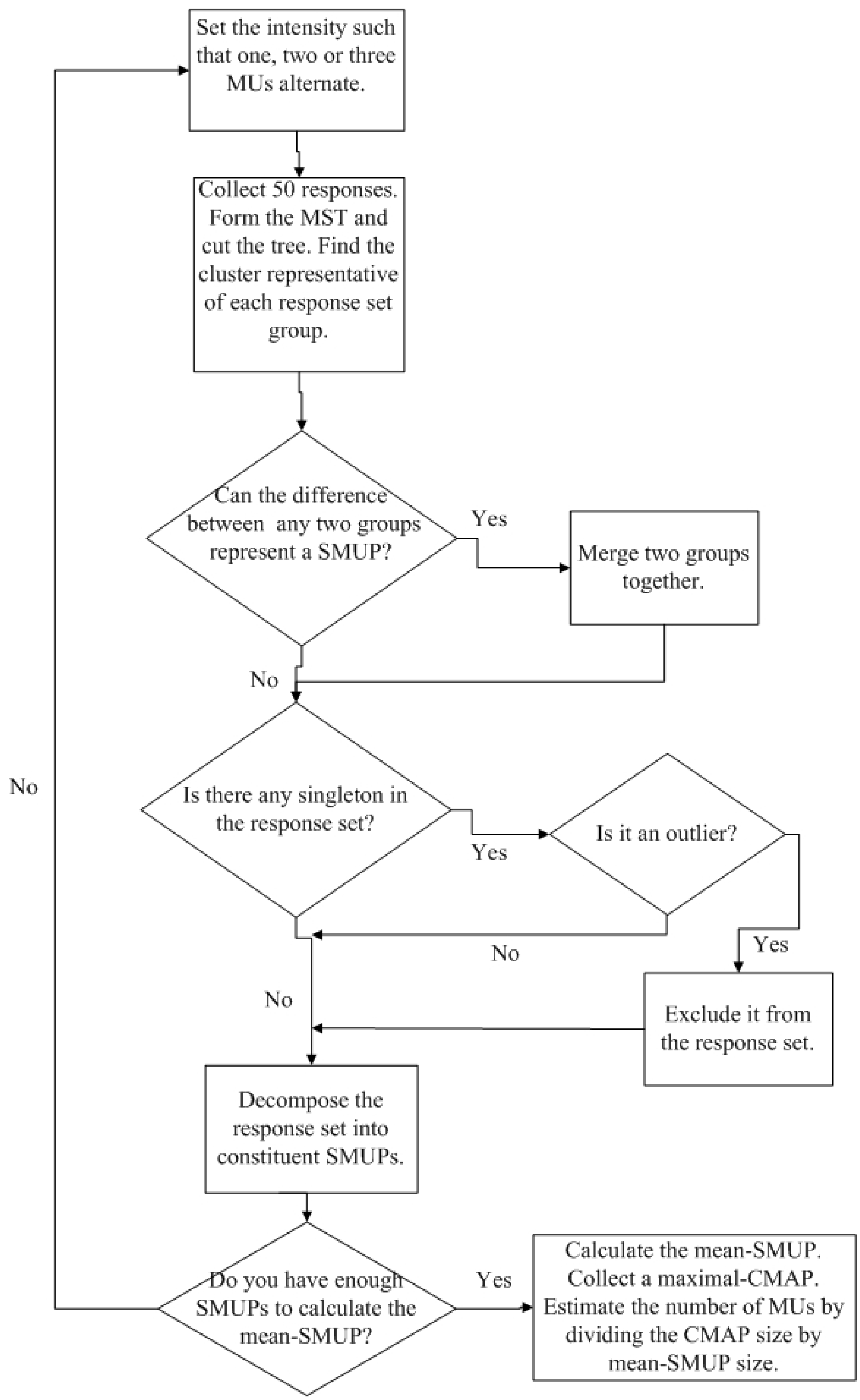


Figure 3.1: Different steps of Automated MPS.

- The MU may never be activated.
- The MU may be activated probabilistically.

At a specific stimulation intensity, no useful information can be obtained from the group of MUs that are always activated, because the number of these MUs is unknown. Likewise, the group of MUs that are never activated are not useful, because they do not affect the acquired potentials. The MUs that are activated probabilistically are the only MUs that can be used for estimating the number of MUs. In general, if the intensity of the train of stimuli is set such that $n$ MUs alternate (i.e., the probability of activation of $n$ MUs is between 0 and 1), $2^n$ different combinations of the alternating MUs might be observed. For example for two alternating MUs, there are four possible combinations. First, none of the MUs are activated. Second, the first MU is activated, but the second MU is not activated. Third, the first MU is not activated, but the second MU is activated. Fourth, both MUs are activated.

In Automated MPS, $N$ stimuli (for example 50 stimuli) are applied to the nerve resulting in the alternation of $n$ MUs and producing $N$ potentials, each of them a possible combination of $n$ SMUPs. These $N$ potentials are detected using surface electrodes. The set including the responses of $n$ alternating MUs to $N$ stimuli is called the response set. Then, the waveforms in the response set are divided to several groups using a hierarchical clustering algorithm (for more details, refer to Section 3.3). If $2^n$ different combinations of the $n$ SMUPs are observed in the response set, it is called a full response set. If some of the combinations are not observed, it is called an incomplete response set.

## 3.3 Clustering the Signals of the Response Set

To divide the $N$ signals of the response set into several groups (up to $2^n$ groups) such that each group represents a unique combination of $n$ SMUPs, a clustering technique should be employed. Assuming that the size and shape of an SMUP does not change in response to repetitive stimulation, the difference between the signals of a group represents the noise. Each response includes one combination of several SMUPs, plus some noise, plus the baseline response. The baseline response is the response of the muscle when none of the $n$ alternating MUs are activated. The baseline response has the lowest energy among the responses, and it is added to all of the responses. A baseline response can be just some noise, or it can be some noise plus the responses of several MUs with the probability of activation of 1.00 at this stimulus intensity (i.e., several MUs that are always activated).

The classification of the response set signals is a challenging task because of the following reasons.

1. Since in Automated MPS, a limited number of responses are collected in the response set (about 50 responses), it is very likely that one or more combinations of $n$ SMUPs are not observed. Hence, although the maximum number of clusters is known, the exact number of clusters is unknown (It can be any number between 1 and $2^n$).

2. Responses collected from the human body always contain noise, the magnitude of which is not constant. Therefore, the signals of each group are not exactly the same. If the noise level is high, it is difficult to discriminate between different groups.

3. Various factors such as patient movement can cause some outliers in the response set that should be detected and excluded. On the other hand, some of the combinations of the SMUPs are observed only once. Hence, all of the singletons cannot be considered as outliers.

### 3.3.1 Clustering Techniques

Clustering techniques sort and subdivide a set of objects in such a way that similar objects fall into the same cluster; whereas dissimilar objects fall into different clusters. Although the term "classification" is used instead of "clustering" loosely, there is a big difference between clustering and classification. In both clustering and classification, an unlabeled object is assigned to the most similar cluster. The difference is in the labels of other objects. In classification, the labels of other objects are known, and the unlabeled object is assigned to one of the pre-determined clusters, whereas in clustering, there is no prior knowledge about the labels of objects. Since there are no pre-determined clusters in Automated MPS, the problem is clustering and not a classification problem.

Data clustering algorithms can be classified into two groups of hierarchical and objective function based (also known as partitional). Hierarchical methods use previously established clusters to find new clusters. They can be agglomerative or divisive. In agglomerative algorithms, each object is considered as a separate cluster at first; then, the clusters are merged together to form larger clusters. Whereas in divisive algorithms, the whole set is considered to be one cluster at first, and then it is divided into smaller clusters [21, 24]. Hierarchical algorithms have two major advantages. First, they can be used to discover clusters with arbitrary shapes. Second, the number of clusters does not need to be predetermined. The major disadvantage of hierarchical techniques is that they are often computationally inefficient [24].

In objective function based algorithms such as the k-means [27], all clusters are determined at once. In these techniques, an initial random or user-defined clustering is performed. Then, the clustering is optimized with respect to some cost criteria. The objective function measures the overall dissimilarity of data objects within each cluster. By minimizing the objective function, the optimal partitioning can be obtained [43]. Objective function based techniques are more efficient computationally and less sensitive to noise compared with hierarchical algorithms. Nevertheless, they have some disadvantages. First, some knowledge about the shape and densities of the clusters is necessary in order to choose the best clustering technique. Second, these methods need the number of clusters in advance. Finding this number is usually challenging and needs a priori knowledge about the data [22, 30].

Since in Automated MPS, a limited number of responses are collected in the response set, and it is very likely that one or more combinations of $n$ SMUPs are not observed, the exact number of clusters is unknown (number of clusters can be any number between 1 and $2^n$). Therefore, objective function based techniques are not suitable candidates for clustering the signals of the response set. The clustering technique used in Automated MPS should be a hierarchical technique.

### 3.3.2 Distance Measure

In order to define a cluster, a measure for similarity or dissimilarity between a pair of objects (pair of responses in the response set) should be determined. Usually, a distance function or metric is used as the measure of dissimilarity. $D(X, Y)$ is a distance function if it is symmetric and always positive. A distance function is a distance metric if it satisfies the triangle inequality and reflexivity conditions [21].

The Euclidian distance is the most popular distance metric used for clustering in the Euclidean space $R^M$. The Euclidian distance is calculated by finding the square of the distance between the coordinates of a pair of objects in each dimension, summing the squares, and finding the square root of that sum. The Euclidian distance between $X =$

$(x_1, x_2, ..., x_M)$ and $Y = (y_1, y_2, ..., y_M)$ is calculated using Equation 3.1.

$$D(X,Y) = \sqrt{\sum_{k=1}^{M} |x_k - y_k|^2} \tag{3.1}$$

where $M$ is the number of dimensions (the number of constituent points of the responses). The Euclidian distance between two points is the distance between two points measured by a ruler.

In Automated MPS, each response is considered as a point in $R^M$ ($M$ is the number of constituent points of each response). Since the Euclidian distance is the most popular distance metric in pattern recognition, and the noise present in surface-detected signals has a Gaussian distribution, the distance between each pair of responses is calculated using the Euclidian distance.

### 3.3.3 Sorting the Response Set

In Automated MPS, a hierarchical clustering technique based on the concept of nearest neighbor is used to group the signals of the response set. Although this concept is widely used in classification, there are some clustering techniques that use it.

Usually, a pattern and its nearest neighbor should have the same labels. A simple non-iterative nearest neighbor clustering technique is proposed by Lu and Fu [25] in which the first object is assigned to the first cluster. Then, the nearest unassigned object to the first pattern is found. If the distance between these two points (the edge weight) is less than a pre-specified threshold, the second pattern is assigned to the same cluster. Otherwise, it is assigned to a new cluster. This procedure is performed until all patterns are labeled. In this algorithm, the number of final partitions depends on the pre-specified threshold. The algorithm finds a large number of clusters if the threshold is small; it finds a small number of clusters if the threshold is large. In Automated MPS, the threshold cannot be pre-specified, because it depends on the noise level and variability of the collected responses (the distance between the two responses of the same group is just noise). Moreover, the result of this algorithm depends on the order of the responses; while Automated MPS needs a clustering method that is independent of the order in which the responses were obtained.

The clustering technique used in the first step of Automated MPS was developed based on the above mentioned algorithm. In Automated MPS, the selection order of the responses is determined based on the nearest neighbor rule, and a tree is built according to this order first. Then, several rules are applied to this tree to cut it to several clusters; the number of clusters depends on these rules (refer to Section 3.3.4).

Any undirected graph consists of a set of nodes and a set of edges (unordered pair of nodes). A tree is a connected graph without any cycle. A spanning tree is a tree that connects all nodes together. A minimum spanning tree is the spanning tree with the minimum weight in compare with other spanning trees.

The algorithm used to build a tree in Automated MPS is very similar to Prim's algorithm for building the minimum spanning tree [32]. The only difference is that the first edge is not selected arbitrarily; it is the edge with the lowest weight (distance). In this tree, each response is a node; each pair of nodes are connected with an edge, and the Euclidian distance between a pair of responses is the edge weight between the two nodes.

### Building the Tree

The process of building this tree is explained in detail below.

- **Step 1.** The distance between each pair of nodes is calculated and a $N$x3 matrix, called "the initial clustering matrix", is defined where $N$ is the number of responses. Each row represents one step of sorting. The third column of each row shows the distance (i.e., the weight) between the signals placed in the first and second columns of that row.

- **Step 2.** Sorting begins by looking for the shortest edge (i.e., the distance between the closest pair of responses). The indices of the corresponding nodes are placed in the first and second column of the first row and the distance between them in the third column.

- **Step 3.** The closest unselected node to the tree is found. The distance between a node and a tree is defined as the distance between the node and the closest node of the tree. To find the closest unselected node, all edges between unselected nodes (i.e., nodes not in the tree) and all nodes of the tree are considered, and the shortest edge is selected. The indices of the nodes of this shortest edge are placed in the first and second column of the next row, and their distance in the third column.

- **Step 4.** Step 3 is repeated until all of the nodes are selected.

In the above procedure, one new response is selected in each row, and the order of selecting the responses shows the order of sorting.

### 3.3.4 Cutting the Tree

After building the tree, the sorted signals of the response set should be clustered, such that each cluster represents a possible combination of the constituent S-MUPs. Since the responses in each group should represent the same combination of SMUPs plus some noise, within-cluster distances are much lower compared to between-cluster distances. Therefore, the built tree is cut by comparing the distances located in the third column of the initial clustering matrix, and looking for any major jump (significant increase) in the values of these distances. The details of this procedure are as follows.

- The overall mean and standard deviation of the distances located in the third column of the initial clustering matrix that represent noise are calculated. To do this, all distances are sorted in an ascending order, and the point at which a distance exceeds 1.5 times the previous distance is found. All distances after this point, which are larger distances, are ignored, and the mean and standard deviation of all distances before this point (smaller distances) are calculated.

- The major jumps in the tree are found, and the tree is cut from those points. For finding a major jump, each distance of the third column of the initial matrix is subtracted from the previous distance, and the resulting value is compared to a threshold calculated according to the following rules, which are determined empirically. If this value is less than the threshold, the response selected in that row is assigned to the same cluster. Otherwise, it is assigned to a new cluster.

  - If the number of signals in the current cluster is less than three, the threshold is five times the overall standard deviation.

  - If the number of signals in the current cluster is more than three, but less than five, the threshold is five times the mean absolute deviation of the distances of the current group (cluster).

  - If the number of signals in the current cluster is more than five, the threshold is five times the standard deviation of the distances of the current group.

The number of clusters in the response set depends on the number of alternating MUs and their probability of activation. One MU can generate one or two groups. Observing three or four groups is the result of the alternation of two MUs. Similarly, three alternating MUs can generate five, six, seven or eight groups in the response set. If there is any outlier group in the response set, the above numbers will be increased.

## 3.4 Finding Cluster Representatives

As mentioned previously, the collected responses of the response set are divided into several groups using a hierarchical clustering algorithm. Since the responses of each group should be compared with those of other groups for further analysis, it is easier to define the cluster representative of each cluster and use it instead of using all the responses of the group. The cluster representative of each group is defined as follows.

- If the number of responses in the group is smaller than four, it is defined as the point-by-point average of all the responses of the group.

- If the number of responses in the group is larger than three, it is defined as the point-by-point average of the first four responses of the group.

Working with the cluster representative of a group is easier than working with all the responses of the group. From now on, comparing two groups will be done by comparing the cluster representatives of the two groups by finding the Euclidian distance between them.

## 3.5 Conclusions

In this chapter, the first step of the Automated MPS was explained. In the first step, a train of stimuli is applied to a motor nerve, activating $n$ MUs probabilistically. One response is detected and stored after each pulse that represents one out of $2^n$ possible combinations of the $n$ SMUPs. The collected signals are sorted and clustered using a nearest neighbor-based hierarchical clustering technique, and the cluster representative of each cluster is calculated. The next chapter explains how the $n$ constituent SMUPs can be extracted using the clusters of the response set.

# Chapter 4

# Decomposition

As mentioned in Chapter 3, in the first step of Automated MPS the waveforms of the response set are clustered, and the cluster representative of each cluster, which should represent one combination of $n$ SMUPs, is calculated. In the second step of Automated MPS, the constituent SMUPs (the responses of individual alternating MUs) need to be recovered.

The decomposition techniques described in this chapter assume that each cluster of the response set represents a unique combination of $n$ SMUPs. This assumption is not true in the following situations. First, it is possible that the clustering algorithm of the first step detects the same combination of SMUPs in two different groups. These groups should be merged together before applying any decomposition technique. Second, several factors such as patient movement cause some outliers in the response set. These outliers do not represent any combination of $n$ SMUPs and should be excluded from the data set before decomposing the responses. The decomposition technique used in the second step depends on the number of response set groups after possible merging and detecting outliers. Figure 4.1 represents different decomposition techniques that Automated MPS uses for different cases. Sections 4.1 and 4.2 describe the merging and excluding steps. Sections 4.3 to 4.5 explain the decomposition procedure for different situations.

## 4.1 Manual or Automatic Merging of Similar Groups

If two groups of the response set represents the same combination of the $n$ SMUPs, they should be merged together before applying any decomposition technique. The merging procedure can be performed manually or automatically. If the merging procedure is performed manually, the operator detects the similar groups visually and decides which groups should be merged together. On the contrary, if the merging procedure is performed automatically, the operator does not play any role in selecting the merging candidates. In this case, the cluster representative of each cluster is compared with the cluster representatives of all other groups. If the difference between the cluster representatives of two groups does not have the physiological properties of a MUP, these groups should be merged together; in this case the difference between the two cluster representatives represents noise. A response cannot represent a MUP in one of the following situations.

1. The peak-to-peak voltage of the response is less than 20 $\mu$v.

2. The area of the response is less than 100 $\mu$v.ms.

## 4.2 Manual Excluding of the Outliers

Several factors such as subject or electrode movement can cause some outliers in the response set. In the current version of Automated MPS, detecting outliers is done manually (visually). The operator detects the outliers by looking at the response set clusters and excludes ones that do not look like a MUP.

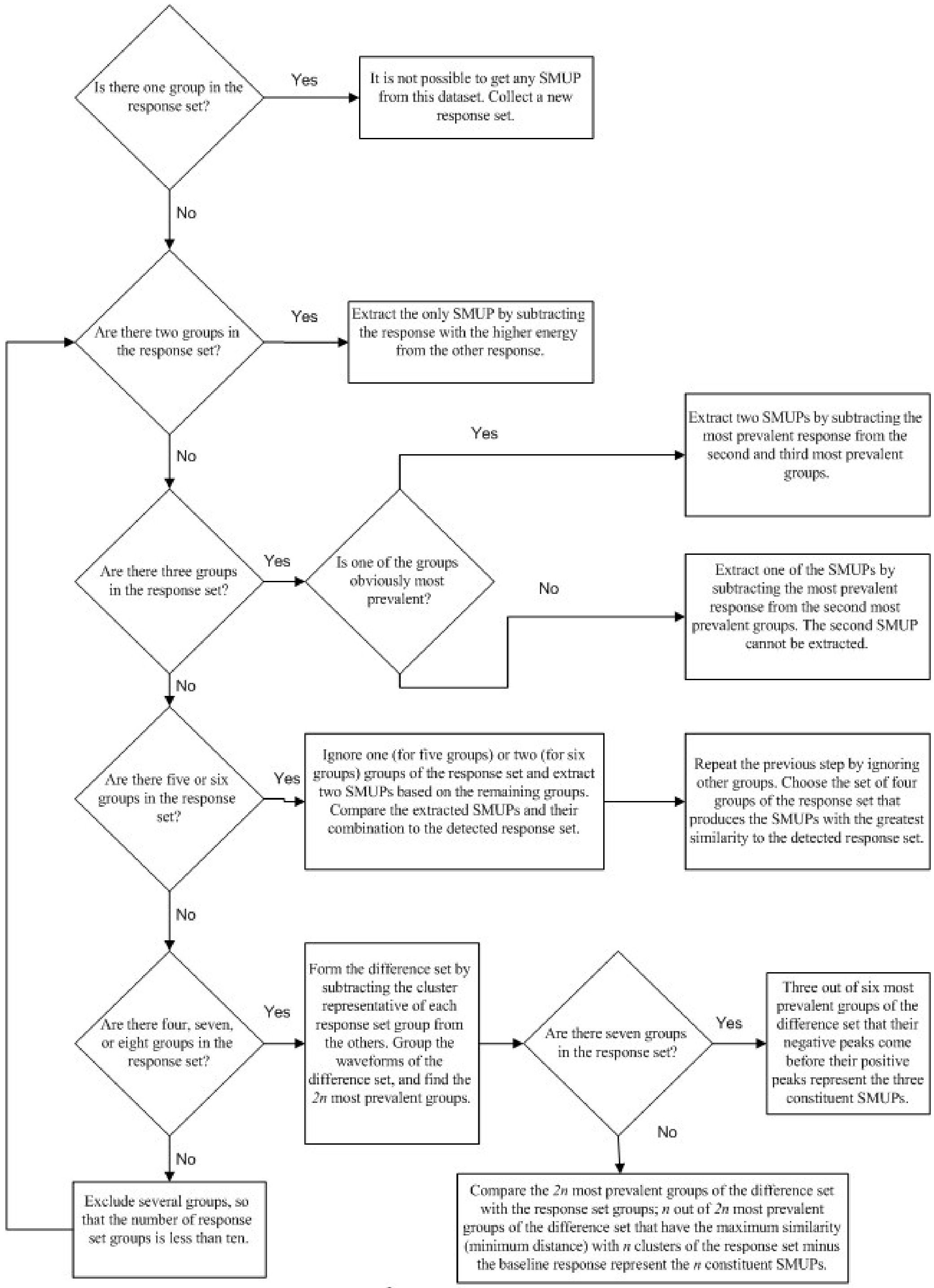


Figure 4.1: Decomposition procedure of Automated MPS depending on the number of response set groups.

## 4.3 Decomposition When One MU Alternates

One MU can respond to a single stimulus in two ways. If it is not activated, the detected response ($R1$) represents the baseline response. If it is activated, the detected response ($R2$) represents one SMUP plus the baseline. If either R1 or R2 are not observed in the response set (i.e., the MU is always activated or always not activated), the response set will be incomplete, and it is not possible to extract an SMUP from the data. If the probability of observing $R1$ after each stimulus is $P$, the probability of observing $R2$ will be $1-P$. For $N$ stimuli, the probabilities of observing only $R1$ or $R2$ are $P^N$ and $(1-P)^N$ respectively. The probability of observing a full response set (observing both $R1$ and $R2$) is $P_{full} = 1 - ((1-P)^N + P^N)$. The probability of observing a full response set increases as the number of stimuli, $N$, increases. For a fixed $N$, this probability is maximum when $P = 0.5$. The same conclusions can be made when the number of alternating MUs is more than one. The maximum probability of observing a full response set is obtained when the probabilities of activation of all alternating MUs are 0.5, and this probability increases as the number of stimuli increases.

For one alternating MU, if a full response set is observed, the only SMUP can be extracted easily by subtracting the smaller response ($R1$, which represents the baseline response) from the larger response ($R2$, which represents the only SMUP plus the baseline response) assuming that there is no phase cancellation in the response set. If more than one MU alternates, the number of response set clusters is more than two, and finding the constituent SMUPs is more complicated. The rest of this chapter explains the details of the Automated MPS decomposition algorithm for cases in which more than one MU are probabilistically activated.

## 4.4 The Number of Stimuli and Alternating MUs

Theoretically, if a sufficient number of stimuli are applied to a motor nerve, all possible combinations of the $n$ alternating SMUPs will be observed. However, considering the factors listed below, the number of stimuli should be limited in practice (In Automated MPS, the number of stimuli is set to 50).

- Automated MPS assumes that the motor nerve system remains stationary during the stimulation (i.e., except for the $n$ alternating MUs, no other MU will be activated). For small numbers of stimuli (around 50 stimuli), this assumption is true. However, this stationarity might not be true for large numbers of stimuli. For example, if the tissue temperature changes slightly during a long period of stimulation, the number of alternating MUs might change.

- The patient should relax and not voluntarily contract the muscle being tested during the course of stimulation. Although this is easy for short periods of stimulation, this might be difficult for long periods of stimulation.

- Although stimulation of a motor nerve is not painful, it is not very comforting. As the frequency and number of stimuli increases, the comfort level of the patient decreases, and anxiety level increases.

Since a limited number of stimuli are used in each sequence of Automated MPS, it is possible that some combinations of alternating SMUPs are not observed in the response set. If the number of alternating MUs is small (i.e., one, two, or three MUs), the probability of observing an incomplete set is not very high. However, this probability increases as the number of alternating MUs increases. If the number of alternating MUs is more than three, it is very likely that several combinations are not observed. Although the decomposition procedure is not difficult if the response set is complete, or if only one of the combinations is absent, it becomes much more complicated if several combinations are absent. Therefore, in Automated MPS, the stimulation intensity is set such that three or fewer MUs are activated probabilistically after each pulse.

## 4.5 Decomposition When More than One MU Alternate

In Section 4.3, the decomposition procedure for cases in which one MU is probabilistically activated was explained. In Section 4.5.1, a new decomposition technique will be introduced for cases in which two or three MUs alternate, and the number of response set clusters is four (full response set, two MUs), seven (incomplete response set, three MUs), or eight (full response set, three MUs). This technique is also able to find the constituent MUs if the number of alternating MUs is more than three, and either a full response set or an incomplete set with one absent combination is observed. However, this technique cannot find the constituent SMUPs if the number of clusters of the response set is three, or if more than one combination of the alternating SMUPs is absent. Section 4.5.2 explains a decomposition solution when there are three groups in the response set, and Section 4.5.3 describes the decomposition procedure for cases in which five or six groups are observed in the response set.

### 4.5.1 Decomposition When the Number of Response Set Clusters Is Four, Seven, or Eight

In this new decomposition technique, the cluster representatives of the response set are used to form a new set, called "the difference set". The waveforms of the difference set are sorted and clustered. Then, the cluster representatives of the $2^n$ most prevalent groups of the difference set are compared with the cluster representatives of the response set minus the baseline response (the response with the lowest energy in the response set) to extract the constituent SMUPs. The details of this algorithm are explained below.

- **Step 1:** The cluster representative of each cluster in the response set is subtracted from the cluster representatives of the other clusters. The resulting waveforms form the difference set.

- **Step 2:** The waveforms of the difference set are sorted using the same algorithm used in the first step of the Automated MPS.

  - **Step 2.1:** The distance between each pair of waveforms is calculated and a $k$x3 matrix, called "the second clustering matrix", is defined where $k$ is the number of waveforms in the difference set. Each row represents one step of sorting. The third column of each row shows the distance between the waveforms placed in the first and second columns of that row.
  - **Step 2.2:** The distance between the closest pair of waveforms is found. The indices of these two waveforms are placed in the first and second column of the first row and the distance between them on the third column.
  - **Step 2.3:** The closets unselected node to the tree (selected nodes) is found. The indices of the corresponding nodes are placed on the first and second column of

the next row, and their distance on the third column.

- **Step 2.4:** Step 2.3 is repeated until all of the waveforms are selected. In each row, one waveform is selected; the order of selecting the waveforms shows the order of sorting.

- **Step 3:** An initial threshold is computed. Since all members of each group are the same response plus some noises, it can be assumed that the threshold is $m$ times (three or four times) the average of the four or five smallest distances of the second clustering matrix. After setting the threshold, the distances of the third column of the second clustering matrix are compared with this threshold. If a distance is smaller than the threshold, the waveform selected in that row is assigned to the same cluster. Otherwise, it is assigned to a new cluster.

- **Step 4:** If the number of clusters is not correct according to Table 4.1, the threshold is changed and step 3 is repeated.

- **Step 5:** The cluster representative of each cluster is calculated.

- **Step 6:** The baseline response (the response with the lowest energy in the response set) is subtracted from the other cluster representatives of the response set. Then, the distance between these waveforms and the cluster representatives of the $2n$ most prevalent groups of the difference set (clusters with the largest number of waveforms) is calculated; $n$ out of $2n$ most prevalent groups of the difference set that have the maximum similarity (minimum distance) with $n$ clusters of the response set minus the baseline represent the $n$ constituent SMUPs. The other $n$ most prevalent clusters represent the inverse of those $n$ SMUPs.

If the number of response set groups is seven, all steps are the same except for Step 6. When one of the combinations is missing, the $n$ constituent SMUPs cannot be obtained by comparing the $2n$ most prevalent groups of the difference set with the response set groups, because it is possible that the absent response is one of the $n$ SMUPs. In response sets with seven groups, assuming that there is no phase cancellation, the $n$ out of $2n$ most prevalent groups of the difference set that have their negative peaks before their positive peaks represent the $n$ SMUPs. The other $n$ groups represent the inverse of these $n$ SMUPs.

Clustering the waveforms of the difference set is easier than clustering the response set, because having the number of response set clusters, the exact number of clusters of the difference set can be specified. For example, if the number of response set clusters is four, there will be 12 waveforms and eight clusters in the difference set. Table 4.1 shows the number of waveforms and clusters of the difference set for different cases, as well as the number of most prevalent groups of the difference set and the number of SMUPs that can be obtained from the data set.

In Step 3 of the above algorithm, if the number of clusters is not correct according to Table 4.1, the threshold should be changed. If the number of clusters is more than what it should be, the threshold should be decreased (. If it is less, the threshold should be increased. Then, the tree should be cut again with this new threshold, and the number of clusters should be calculated. The threshold should be changed until the correct number of clusters is obtained. If the correct number of clusters is not obtained after 30 times, the process is terminated, and the $2n$ most prevalent groups are selected.

### Justification

Suppose that $S1$ and $S2$ represent the responses of two MUs. If these two MUs alternate, they can produce up to four combinations:

Table 4.1: Properties of the response and difference sets for different cases.

| Num. Of Response set Clusters | Num. of Diff. Set Waveforms | Num. of Diff. Set Clusters | Num. of Most Prevalent Groups | Num. of SMUPs |
|---|---|---|---|---|
| 4 | 12 | 8 | 4 | 2 |
| 7 | 42 | 26 | 6 | 3 |
| 8 | 56 | 24 | 6 | 3 |

- $R1$=$S1$+baseline (only the first MU is activated).
- $R2$=$S2$+baseline (only the second MU is activated).
- $R3$=$S1$+$S2$+baseline (both MUs are activated).
- $R4$=baseline (none of the MUs are activated).

Assuming that all of the four combinations are observed in the response set, there will be 12 waveforms in the difference set. If these 12 waveforms are clustered, there will be eight groups, four of them have two members representing $S1$, $-S1$, $S2$ and $-S2$, and four of them have one member representing $S1 + S2$, $-(S1 + S2)$, $S1 - S2$ and $S2 - S1$. As mentioned previously, the cluster representatives of the four most prevalent groups represent the two SMUPs and their inverses. If the cluster representatives of these four groups are compared with the cluster representatives of the responses minus the baseline response, $S1$ and $S2$ are extracted. Two of these four cluster representatives, S1 and S2, have the maximum similarity to two cluster representatives of the response set minus the baseline response ($R1$-baseline and $R2$-baseline). Table 4.2 explains the decomposition procedure for the above cases.

Table 4.2: Sorting and decomposition procedures when two MUs alternate and a full response set is observed ($S_i$=$i$-th SMUP, and $R_i$=$i$-th response set cluster representative).

| Constituent SMUPs | Response Set Clusters | Diff. Set Waveforms | Clusters of the Diff. Set | Num. Of Waveforms in Each Cluster |
|---|---|---|---|---|
| S1 | R1=S1+baseline | R1-R2=S1-S2 | S1 | 2 |
| S2 | R2=S2+baseline | R1-R3=-S2 | S2 | 2 |
| | R3=S1+S2+baseline | R1-R4=S1 | - S1 | 2 |
| | R4=baseline | R2-R1=S2-S1 | -S2 | 2 |
| | | R2-R3=-S1 | S1+S2 | 1 |
| | | R2-R4=S2 | -(S1+S2) | 1 |
| | | R3-R1=S1 | S1-S2 | 1 |
| | | R3-R2=S2 | -(S1-S2) | 1 |
| | | R3-R4=S1+S2 | | |
| | | R4-R1=-S1 | | |
| | | R4-R2=-S2 | | |
| | | R4-R3=-S1-S2 | | |

If the number of alternating MUs is three, and a full response set is observed, the most prevalent groups are six groups, each of them having four waveforms representing the three constituent SMUPs and their inverses. Similarly, if three MUs alternate, and one of the combinations of the SMUPs is absent in the response set (it does not matter which combination is absent), there will be six most prevalent groups in the difference set, each of them having three waveforms. These six groups represent the three alternating SMUPs and their inverses. Tables 4.3 and 4.4 describe the decomposition procedure for these two cases.

The above decomposition technique cannot find all constituent SMUPs if the number of response set groups is three, five, or six. In these situations, the $2n$ most prevalent groups of the difference set do not represent the $n$ SMUPs and their inverses. If the number of response set groups is three, the decomposition technique used in MUESA method can be used for extracting the SMUPs [34]. However, if there are five or six groups, none of the current decomposition techniques can find all three constituent SMUPs. In these cases, only two SMUPs can be obtained from the data set. Sections 4.5.2 and 4.5.3 describe the decomposition procedure for these cases.

### 4.5.2 Decomposition When the Number of Response Set Clusters is Three

If the number of response set clusters is three, one or two SMUPs can be obtained from the data set using the decomposition technique of the MUESA method [34]. This decomposition technique is based on the number of responses in the response set clusters. If one of the response set groups is obviously more prevalent than the other two groups, two constituent SMUPs can be calculated by subtracting the cluster representatives of the second and third most prevalent groups from the cluster representative of the most prevalent

Table 4.3: Sorting and decomposition procedures when three MUs alternate and a full response set is observed ($S_i$=$i$-th SMUP, and $R_i$=$i$-th response set cluster representative).

| Constituent SMUPs | Response Set Clusters | Diff. Set Waveforms | Diff. Set Waveforms | Clusters of the Diff. Set | Num. Of Waveforms in Each Cluster |
|---|---|---|---|---|---|
| S1 | R1=S1+b | R1-R2=S1-S2 | R5-R1=S2 | S1 | 4 |
| S2 | R2= S2+b | R1-R3=S1-S3 | R5-R2=S1 | S2 | 4 |
| S3 | R3=S3+b | R1-R4=S1 | R5-R3=S1+S2-S3 | S3 | 4 |
| | R4=b | R1-R5=-S2 | R5-R4=S1+S2 | -S1 | 4 |
| | R5=S1+S2+b | R1-R6=-S3 | R5-R6=S2-S3 | -S2 | 4 |
| | R6=S1+S3+b | R1-R7=S1-S2-S3 | R5-R7=S1-S3 | -S3 | 4 |
| | R7=S2+S3+b | R1-R8=-(S2+S3) | R5-R8=-S3 | S1-S2 | 2 |
| | R8=S1+S2+S3+b | R2-R1=S2-S1 | R6-R1=S3 | -(S1-S2) | 2 |
| | | R2-R3=S2-S3 | R6-R2=S1+S3-S2 | S1-S3 | 2 |
| | | R2-R4=S2 | R6-R3=S1 | -(S1-S3) | 2 |
| | | R2-R5=-S1 | R6-R4=S1+S3 | S2-S3 | 2 |
| | | R2-R6=S2-S1-S3 | R6-R5=S3-S2 | -(S2-S3) | 2 |
| | | R2-R7=-S3 | R6-R7=S1-S2 | S1+S2 | 2 |
| | | R2-R8=-S1-S3 | R6-R8=-S2 | -(S1+S2) | 2 |
| | | R3-R1=S3-S1 | R7-R1=S2+S3-S1 | S1+S3 | 2 |
| | | R3-R2=S3-S2 | R7-R2=S3 | -(S1+S3) | 2 |
| | | R3-R4=S3 | R7-R3=S2 | S2+S3 | 2 |
| | | R3-R5=S3-S1-S2 | R7-R4=S2+S3 | -(S2+S3) | 2 |
| | | R3-R6=-S1 | R7-R5=S3-S1 | S1-S2-S3 | 1 |
| | | R3-R7=-S2 | R7-R6=S2-S1 | -(S1-S2-S3) | 1 |
| | | R3-R8=-S1-S2 | R7-R8=-S1 | S2-S1-S3 | 1 |
| | | R4-R1=-S1 | R8-R1=S2+S3 | -(S2-S1-S3) | 1 |
| | | R4-R2=-S2 | R8-R2=S1+S3 | S3-S1-S2 | 1 |
| | | R4-R3=-S3 | R8-R3=S1+S2 | -(S3-S1-S2) | 1 |
| | | R4-R5=-S1-S2 | R8-R4=S1+S2+S3 | S1+S2+S3 | 1 |
| | | R4-R6=-S1-S3 | R8-R5=S3 | -(S1+S2+S3) | 1 |
| | | R4-R7=-S2-S3 | R8-R6=S2 | | |
| | | R4-R8=-S1-S2-S3 | R8-R7=S1 | | |

Table 4.4: Sorting and decomposition procedures when three MUs alternate and one of the combinations (S1+S2+S3+baseline) is absent ($S_i$=$i$-th SMUP, and $R_i$=$i$-th response set cluster representative).

| Constituent SMUPs | Response Set Clusters | Diff. Set Waveforms | Diff. Set Waveforms | Clusters of the Diff. Set | Num. Of Waveforms in Each Cluster |
|---|---|---|---|---|---|
| S1 | R1=S1+b | R1-R2=S1-S2 | R5-R1=S2 | S1 | 3 |
| S2 | R2= S2+b | R1-R3=S1-S3 | R5-R2=S1 | S2 | 3 |
| S3 | R3=S3+b | R1-R4=S1 | R5-R3=S1+S2-S3 | S3 | 3 |
| | R4=b | R1-R5=-S2 | R5-R4=S1+S2 | -S1 | 3 |
| | R5=S1+S2+b | R1-R6=-S3 | R5-R6=S2-S3 | -S2 | 3 |
| | R6=S1+S3+b | R1-R7=S1-S2-S3 | R5-R7=S1-S3 | -S3 | 3 |
| | R7=S2+S3+b | R2-R1=S2-S1 | R6-R1=S3 | S1-S2 | 2 |
| | | R2-R3=S2-S3 | R6-R2=S1+S3-S2 | -(S1-S2) | 2 |
| | | R2-R4=S2 | R6-R3=S1 | S1-S3 | 2 |
| | | R2-R5=-S1 | R6-R4=S1+S3 | -(S1-S3) | 2 |
| | | R2-R6=S2-S1-S3 | R6-R5=S3-S2 | S2-S3 | 2 |
| | | R2-R7=-S3 | R6-R7=S1-S2 | -(S2-S3) | 2 |
| | | R3-R1=S3-S1 | R7-R1=S2+S3-S1 | S1+S2 | 1 |
| | | R3-R2=S3-S2 | R7-R2=S3 | -(S1+S2) | 1 |
| | | R3-R4=S3 | R7-R3=S2 | S1+S3 | 1 |
| | | R3-R5=S3-S1-S2 | R7-R4=S2+S3 | -(S1+S3) | 1 |
| | | R3-R6=-S1 | R7-R5=S3-S1 | S2+S3 | 1 |
| | | R3-R7=-S2 | R7-R6=S2-S1 | -(S2+S3) | 1 |
| | | R4-R1=-S1 | | S1-S2-S3 | 1 |
| | | R4-R2=-S2 | | -(S1-S2-S3) | 1 |
| | | R4-R3=-S3 | | S2-S1-S3 | 1 |
| | | R4-R5=-S1-S2 | | -(S2-S1-S3) | 1 |
| | | R4-R6=-S1-S3 | | S3-S1-S2 | 1 |
| | | R4-R7=-S2-S3 | | -(S3-S1-S2) | 1 |

groups. On the other hand, if there are two most prevalent groups with a relatively close number of responses, only one SMUP can be obtained from the data set by subtracting the cluster representative of the second most prevalent group from that of the most prevalent group.

### 4.5.3 Decomposition When the Number of Response Set Clusters Is Five or Six

If the number of response set clusters is five, only two SMUPs can be obtained from the data set. To extract these two SMUPs, the first clusters is ignored, and two SMUPs are calculated by applying the decomposition technique described in Section 4.5.1 to the four remaining groups. Then, the second cluster is ignored, and the SMUPs are calculated using the other four groups. The same calculations are performed, each time ignoring one group and extracting two SMUPs. Since there are five groups and one group is ignored each time, five possible cases should be studied. Among these five cases, the one that produces the response set which is most similar to the original response set should be chosen. To do this, a response set (three combinations, S1, S2 and S1+S2) is generated for each case using the two extracted SMUPs. Then, these generated response sets are compared to the detected response set, and the one with the greatest similarity to the original response set is selected. The extracted SMUPs of that case are the SMUPs that can be obtained from this data set.

If the number of response set clusters is six, the decomposition is almost the same. The only difference is that each time, two groups are ignored, and the SMUPs are calculated using the four remaining groups. Since the number of response set clusters is six, and two groups are ignored each time, C(6,2)=15 possible cases should be studied.

## 4.6 Summary

In this chapter, the decomposition step of the Automated MPS technique was described. Due to several factors, the number of stimuli should be limited to a number of the order of 50. With a limited number of stimuli, the probability of not observing several combinations of the SMUPs becomes high if the number of alternating MUs is more than two or three. Since the decomposition procedure becomes difficult if more than one combination is absent, in Automated MPS, each train contains 50 stimuli, and the stimulus intensity is set such that after each pulse, only one, two or three MUs are probabilistically activated. Before applying a decomposition technique, the similar groups of the response set are merged together, and the outliers are excluded from the data set. A new decomposition technique is introduced in Section 4.5.1, which can handle most of the cases. A summary of the decomposition solution for each of the possible cases is explained below (Refer to Figure 4.1).

- **Case 1:** If the number of response set groups is one, no SMUP can be obtained.
- **Case 2:** If the number of response set groups is two, the only SMUP can be obtained by subtracting the smaller response from the larger response.
- **Case 3:** If the number of response set groups is three, one or two SMUPs can be obtained using the decomposition technique of the MUESA technique.
- **Case 4:** If the number of response set groups is four, seven, or eight, all constituent SMUPs can be obtained using the decomposition technique introduced in Section 4.5.1.
- **Case 5:** If the number of response set groups is five or six, two SMUPs can be obtained from the data set by excluding one or two groups of the response set, and

extracting the SMUPs using the four remaining groups.

# Chapter 5

# Experiments and Results

Automated MPS is a new tool for decomposing a CMAP into its constituent SMUPs. These SMUPs later can be used to calculate an unbiased mean-SMUP and estimate the number of MUs of the muscle. Automated MPS finds the constituent SMUPs while several MUs alternate. The details of Automated MPS were described in Chapters 3 and 4. This chapter describes experiments performed to evaluate the ability of Automated MPS to correctly identify the constituent SMUPs when different combinations of the responses of $n$ alternating MUs are observed in a response set.

## 5.1 Implementation

Automated MPS has been implemented in C++, and the hardware used for stimulating the muscle and recording the resulting signals is Comperio Hardware. Automated MPS has been added to MPS. Therefore, the operator working with Automated MPS can use Automated MPS, as well as MPS for collecting single SMUPs. MPS and Automated MPS share some parts including acquiring a maximal CMAP, aligning the single SMUPs to calculate the mean-SMUP, and estimating the number of MUs by dividing the CMAP size by the mean-SMUP size.

## 5.2 Parameter Setting

The low-pass and high-pass frequencies are set to 5 and 5000 Hz, covering the typical range of frequencies observed in surface detected EMG signals. The sampling frequency is set to 31.25 kHz, and there is a one second time interval between stimuli. After each stimulus, 200 ms of the detected signal is recorded (Usually the first 50ms after each pulse contains valuable information); each recorded response consists of 6250 data points.

## 5.3 Data Generation

To evaluate the performance of the first step of Automated MPS, several experiments were performed using simulated response sets, as well as 10 response sets collected from patients. Most of the experiments used simulated response sets, because the number of alternating MUs in simulated response sets is known and does not need to be determined visually, as opposed to real response sets in which there are no gold standards. For evaluating the performance of the second step, the outputs of the first step (i.e., the response set clusters) are used as the inputs of the second step for both simulated and real data.

According to Section 4.4, considering the following factors, the number of responses in each response set is limited to 50. First, the patient should be comfortable and relaxed; he/she should not do any voluntary contraction. Second, the electrodes should not move during collection of the responses. Third, the size of the same combination of several SMUPs should not change with repetitive stimuli. On the other hand, since the possibility of observing an incomplete response set (not observing some combinations) increases as the number of alternating MUs increases, the number of alternating MUs is limited to three.

### 5.3.1 Simulated data

This section describes how the waveforms of each simulated response set were generated. To generate each response set, one, two or three SMUPs were selected from the SMUPs of a data set collected with MPS; then, a probability of firing was assigned to each selected SMUP. Moreover, one of the collected waveforms was selected as the baseline response with a probability of activation of one; this waveform was just noise or a waveform having small energy compared with other selected SMUPs. Table 5.1 shows the properties of SMUPs selected from 32 data sets collected with MPS, which are used to generate the simulated response sets. This table shows the probability of firing assigned to each selected SMUP, as well as the probability of observing each combination of the selected SMUPs, which is calculated by multiplying the probability of firing of the fired MUs by the probability of not firing of the MUs which are not in that combination. For example, the probability of observing $R3$ (the third SMUP plus baseline) is calculated as follows. $P(R3 = S3 + baseline) = (1 - P(S1)).(1 - P(S2)).P(S3)$

To generate 50 responses of each response set, 50 random numbers between 0 and 1 were generated. After generating each random number, some of the $n$ selected MUs were fired based on their probability of activation. If the probability of activation of a MU was less than the generated random number, it was fired; a MU was not fired if its probability of activation was higher than the generated random number. The SMUPs of the fired MUs were added together, and the baseline response was added to them (The probability of firing the baseline response is always equal to one). If no noise was added to these responses, the performance of the Automated MPS in decomposing the response set would be 100%, because all the waveforms of each response set cluster were exactly the same. For real data, the presence of noise makes the clustering challenging. To simulate this, noise was added to each response; this noise had a Gaussian distribution with zero mean and maximum peak to peak amplitude of 10 $\mu$V. The generated response set included 50 responses, each of them a combination of $n$ SMUPs plus some noise. Automated MPS was applied to this response set trying to recover the constituent SMUPs.

To generate a full response set (two, four or eight response set groups), 50 random numbers were generated. If all possible combinations of the $n$ selected SMUPs were not observed in the response set, the response set was ignored, and new sets of 50 random numbers were generated until the generated response set contained all $2^n$ possible combinations of the $n$ selected SMUPs.

On the other hand, an incomplete set (three, five, six or seven response set groups) was generated when either some of the combinations of the $n$ SMUPs were not observed, because their probabilities of observation were so small, or when the data was simulated such that specific combinations of the $n$ SMUPs were not generated. The former was used to generate response sets with three clusters, while the latter was used to generate response sets with five, six, or seven clusters.

### 5.3.2 Real Data

Other than the simulated response sets explained above, Automated MPS was applied to 10 response sets collected from stimulating the median nerve of control subjects in our lab.

## 5.4 Selecting a Window for Distance Calculation

Using the whole 200ms recorded waveform in the calculation of the distance between two waveforms has a noticeable negative effect on the performance of Automated MPS. Most of the collected responses have a large and sharp peak at the beginning, called the stimulus artifact, which is the result of the stimulus and is not a biological response. The actual biological response follows this stimulus artifact after a few milliseconds. Including the stimulus artifact in the calculation of the distances between each pair of waveforms can disturb the results of clustering the responses. On the other hand, some of the responses contain an F-response after a few milliseconds from the main response. The F-response or any other major change that happens after approximately 10ms from the negative peak of each response should not be included in distance calculations. Therefore, a window containing only the significant parts of the collected responses is selected for calculating the distance between the waveforms. The significant part of an SMUP or a combination of several SMUPs usually begins at the onset of the waveform and extends to either its end point or 10ms after its negative peak. In our experiments, the start point of the selected window is set to the median trimmed average onset of all responses. On the other hand, since small responses such as the baseline response might have very late end points, they should be ignored in the calculation of the end point of the selected window. Therefore, the end point is set as median trimmed peak of the half largest responses of the response set plus 10ms.

Table 5.1: Properties of the data sets used to generate simulated response sets, the probability of firing $i-th$ SMUP($Si$) and observing the $i-th$ combination of the responses($Ri$)

| **Data set** | **Num. of selected SMUPs** | **prob. of firing S1** | **prob. of firing S2** | **prob. of firing S3** | **Prob. of observing R1** | **Prob. of observing R2** | **Prob. of observing R3** | **Prob. of observing R4** | **Prob. of observing R5** | **Prob. of observing R6** | **Prob. of observing R7** | **Prob. of observing R8** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 3 | 0.7 | 0.3 | 0.8 | 0.042 | 0.098 | 0.018 | 0.042 | 0.168 | 0.392 | 0.072 | 0.168 |
| **2** | 2 | 0.4 | 0.7 | - | 0.18 | 0.12 | 0.42 | 0.28 | - | - | - | - |
| **3** | 2 | 0.8 | 0.6 | - | 0.08 | 0.32 | 0.12 | 0.48 | - | - | - | - |
| **4** | 3 | 0.6 | 0.4 | 0.7 | 0.072 | 0.108 | 0.048 | 0.072 | 0.168 | 0.252 | 0.112 | 0.168 |
| **5** | 2 | 0.8 | 0.4 | - | 0.12 | 0.48 | 0.08 | 0.32 | - | - | - | - |
| **6** | 2 | 0.6 | 0.8 | - | 0.08 | 0.12 | 0.32 | 0.48 | - | - | - | - |
| **7** | 3 | 0.7 | 0.4 | 0.3 | 0.126 | 0.294 | 0.084 | 0.196 | 0.054 | 0.126 | 0.036 | 0.084 |
| **8** | 3 | 0.6 | 0.3 | 0.5 | 0.14 | 0.21 | 0.06 | 0.09 | 0.14 | 0.21 | 0.06 | 0.09 |
| **9** | 2 | 0.2 | 0.7 | - | 0.24 | 0.06 | 0.56 | 0.14 | - | - | - | - |
| **10** | 3 | 0.3 | 0.2 | 0.7 | 0.168 | 0.072 | 0.042 | 0.018 | 0.392 | 0.168 | 0.098 | 0.042 |
| **11** | 2 | 0.3 | 0.2 | - | 0.56 | 0.24 | 0.14 | 0.06 | - | - | - | - |
| **12** | 3 | 0.3 | 0.2 | 0.5 | 0.28 | 0.12 | 0.07 | 0.03 | 0.28 | 0.12 | 0.07 | 0.03 |
| **13** | 3 | 0.3 | 0.2 | 0.7 | 0.168 | 0.072 | 0.042 | 0.018 | 0.392 | 0.168 | 0.098 | 0.042 |
| **14** | 2 | 0.4 | 0.3 | - | 0.42 | 0.28 | 0.18 | 0.12 | - | - | - | - |
| **15** | 3 | 0.3 | 0.4 | 0.3 | 0.294 | 0.126 | 0.196 | 0.084 | 0.126 | 0.054 | 0.084 | 0.036 |
| **16** | 3 | 0.2 | 0.2 | 0.5 | 0.32 | 0.08 | 0.08 | 0.02 | 0.32 | 0.08 | 0.08 | 0.02 |
| **17** | 2 | 0.3 | 0.1 | - | 0.63 | 0.27 | 0.07 | 0.03 | - | - | - | - |
| **18** | 3 | 0.5 | 0.5 | 0.5 | 0.125 | 0.125 | 0.125 | 0.125 | 0.125 | 0.125 | 0.125 | 0.125 |
| **19** | 2 | 0.6 | 0.4 | - | 0.24 | 0.36 | 0.16 | 0.24 | - | - | - | - |
| **20** | 3 | 0.4 | 0.5 | 0.3 | 0.21 | 0.14 | 0.21 | 0.14 | 0.09 | 0.06 | 0.09 | 0.06 |
| **21** | 2 | 0.3 | 0.2 | - | 0.56 | 0.24 | 0.14 | 0.06 | - | - | - | |
| **22** | 2 | 0.1 | 0.7 | - | 0.27 | 0.03 | 0.63 | 0.07 | - | - | - | |
| **23** | 2 | 0.1 | 0.95 | - | 0.045 | 0.005 | 0.855 | 0.095 | - | - | - | |
| **24** | 2 | 0.9 | 0.1 | - | 0.09 | 0.81 | 0.01 | 0.09 | - | - | - | |
| **25** | 2 | 0.9 | 0.1 | - | 0.09 | 0.81 | 0.01 | 0.09 | - | - | - | |
| **26** | 2 | 0.9 | 0.1 | - | 0.09 | 0.81 | 0.01 | 0.09 | - | - | - | |
| **27** | 2 | 0.9 | 0.1 | - | 0.09 | 0.81 | 0.01 | 0.09 | - | - | - | |
| **28** | 2 | 0.95 | 0.1 | - | 0.045 | 0.855 | 0.005 | 0.095 | - | - | - | |
| **29** | 2 | 0.1 | 0.95 | - | 0.045 | 0.005 | 0.855 | 0.095 | - | - | - | |
| **30** | 2 | 0.1 | 0.95 | - | 0.045 | 0.005 | 0.855 | 0.095 | - | - | - | |
| **31** | 2 | 0.85 | 0.05 | - | 0.1425 | 0.8075 | 0.0075 | 0.0425 | - | - | - | |
| **32** | 2 | 0.85 | 0.05 | - | 0.1425 | 0.8075 | 0.0075 | 0.0425 | - | - | - | |

## 5.5 Results

### 5.5.1 Full Response Sets

Automated MPS was applied to 60 simulated full response sets (30 response sets with four groups, and 30 response sets with eight groups). To generate these 60 response sets, 20 data sets were selected from the data sets of Table 5.1; 10 of them having two SMUPs and 10 of them having three SMUPs. Then, three full response sets were generated using the selected SMUPs of each data set.

Table 5.2 shows the results of applying Automated MPS to these response sets. The results of each experiment are shown with five numbers representing the number of detected response set groups, the number of SMUPs detected correctly, the number of responses assigned to a wrong cluster, the number of most prevalent groups detected incorrectly, and the performance of Automated MPS respectively. Table 5.2 shows that Automated MPS classified 1481 out of 1500 responses (98.7%) and found 147 out of 150 SMUPs (97.5%) correctly. It could not find one of the SMUPs in three experiments (Exp. 19-2, 19-3 and 29-1), because in those experiments the SMUP sizes were so small, such that after adding noise, differentiating between certain clusters was difficult, even visually. Most of the misclassifications (18 out of 19) also happened in these three experiments. From the above results, it can be concluded that Automated MPS works very well if the observed response set is full, the selected window is calculated properly, and the noise level is not high compared with SMUPs sizes.

### 5.5.2 Response Sets with Seven Groups

To generate a response set with seven groups, the data sets were simulated such that one combination of the three selected SMUPs was not generated in the response set. Ten

Table 5.2: Results of applying Automated MPS to full response sets.

| | | | **Experiment 1** | | | | | **Experiment 2** | | | | | **Experiment 3** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Data set** | **Num. of SMUPs** | **Num. of response set groups** | **Num. of detected response set groups** | **Num. of correctly detected SMUPs** | **Num. of misclassified responses** | **Num. of wrong most prevalent groups** | **Performance in finding the SMUPs** | **Num. of detected response set groups** | **Num. of correctly detected SMUPs** | **Num. of misclassified responses** | **Num. of wrong most prevalent groups** | **Performance in finding the SMUPs** | **Num. of detected response set groups** | **Num. of correctly detected SMUPs** | **Num. of misclassified responses** | **Num. of wrong most prevalent groups** | **Performance in finding the SMUPs** |
| **1** | 3 | 8 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 |
| **4** | 3 | 8 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 |
| **7** | 3 | 8 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 |
| **8** | 3 | 8 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 |
| **10** | 3 | 8 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 |
| **13** | 3 | 8 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 | 9 | 3 | 0 | 0 | 100 |
| **15** | 3 | 8 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 |
| **16** | 3 | 8 | 8 | 3 | 0 | 0 | 100 | 9 | 3 | 1 | 0 | 100 | 8 | 3 | 0 | 0 | 100 |
| **18** | 3 | 8 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 1 | 100 | 8 | 3 | 0 | 0 | 100 |
| **20** | 3 | 8 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 |
| **3** | 2 | 4 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 |
| **6** | 2 | 4 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 |
| **9** | 2 | 4 | 4 | 2 | 0 | 1 | 100 | 4 | 2 | 0 | 1 | 100 | 4 | 2 | 0 | 1 | 100 |
| **11** | 2 | 4 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 |
| **14** | 2 | 4 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 |
| **17** | 2 | 4 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 |
| **19** | 2 | 4 | 4 | 2 | 0 | 0 | 100 | 3 | 1 | 10 | 0 | 50 | 3 | 1 | 8 | 0 | 50 |
| **22** | 2 | 4 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 |
| **23** | 2 | 4 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 | 4 | 2 | 0 | 0 | 100 |
| **29** | 2 | 4 | 4 | 1 | 0 | 3 | 50 | 4 | 2 | 0 | 2 | 100 | 4 | 2 | 0 | 1 | 100 |
| **Total** | | **120** | **120** | **49** | **0** | **4** | **-** | **120** | **49** | **11** | **4** | **-** | **120** | **49** | **8** | **2** | **-** |
| | | | **Performance = 97.5%** | | | | | **Performance = 97.5%** | | | | | **Performance = 97.5%** | | | | |

data sets with three SMUPs were selected from Table 5.1, and three response sets were generated based on each data set. In the first two experiments of each data set, the absent combination was either $S1 + S2 + baseline$ or $S1 + S2 + S3 + baseline$; in the third experiment, the absent response was the baseline response.

The results of applying Automated MPS to the above 30 response sets are shown in Table 5.3. Automated MPS classified 1490 out of 1500 responses (99.3%) correctly. It found 27 out of 30 SMUPs (90%) correctly in the first and third set of experiments, and 28 out of 30 SMUPs (93.3%) in the second sets of experiments. Selecting a bad window for distance calculations was the reason for extracting a wrong SMUP in Exp. 7-1, 7-2, 16-1 and 16-3. In these experiments, the majority of responses have late onsets. Therefore, choosing the median trimmed average onset as the start point of the selected window (Refer to Section 5.4) lead to loosing some significant parts of the responses with early onsets and detecting the inverse of the third SMUP as an SMUP. On the other hand, in Exp. 10-1, 10-2, 10-3 and 20-3, one SMUP was detected incorrectly, because one of the selected SMUPs was very small, while the other one was very large, such that it was difficult to differentiate between the large SMUP and the combination of these two SMUPs.

### 5.5.3 Response Sets with Three Groups

To generate response sets with three groups, 10 data sets with two SMUPs were selected from Table 5.1. These data sets are the ones in which the probability of observing one of the combinations of the two SMUPs is very small (0.01 or less). Therefore, it is very likely that the generated response set does not contain this combination, and therefore has three clusters. For each data set, three experiments were performed with three different response sets.

Table 5.4 shows the results of these 30 experiments. Automated MPS classified 1489

Table 5.3: Results of applying Automated MPS to incomplete response sets with 7 groups (Three MUs alternate and one group is absent).

| | | | **Experiment 1** | | | | | **Experiment 2** | | | | | **Experiment 3** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Data set** | **Num. of SMUPs** | **Num. of response set groups** | **Num. of detected response set groups** | **Num. of correctly detected SMUPs** | **Num. of misclassified responses** | **Num. of wrong most prevalent groups** | **Performance in finding the SMUPs** | **Num. of detected response set groups** | **Num. of correctly detected SMUPs** | **Num. of misclassified responses** | **Num. of wrong most prevalent groups** | **Performance in finding the SMUPs** | **Num. of detected response set groups** | **Num. of correctly detected SMUPs** | **Num. of misclassified responses** | **Num. of wrong most prevalent groups** | **Performance in finding the SMUPs** |
| **1** | 3 | 7 | 8 | 3 | 0 | 1 | 100 | 7 | 3 | 0 | 0 | 100 | 7 | 3 | 0 | 0 | 100 |
| **7** | 3 | 7 | 7 | 2 | 0 | 0 | 66.6 | 7 | 2 | 0 | 0 | 66.6 | 7 | 3 | 0 | 0 | 100 |
| **8** | 3 | 7 | 7 | 3 | 0 | 0 | 100 | 9 | 3 | 0 | 2 | 100 | 8 | 3 | 0 | 0 | 100 |
| **10** | 3 | 7 | 7 | 2 | 0 | 2 | 66.6 | 7 | 2 | 0 | 2 | 66.6 | 7 | 2 | 0 | 2 | 66.6 |
| **12** | 3 | 7 | 7 | 3 | 0 | 0 | 100 | 7 | 3 | 0 | 0 | 100 | 7 | 3 | 0 | 0 | 100 |
| **13** | 3 | 7 | 7 | 3 | 0 | 0 | 100 | 8 | 3 | 0 | 0 | 100 | 7 | 3 | 0 | 0 | 100 |
| **15** | 3 | 7 | 7 | 3 | 1 | 0 | 100 | 8 | 3 | 1 | 1 | 100 | 7 | 3 | 0 | 0 | 100 |
| **16** | 3 | 7 | 7 | 2 | 1 | 0 | 66.6 | 7 | 3 | 0 | 0 | 100 | 7 | 2 | 2 | 0 | 66.6 |
| **18** | 3 | 7 | 7 | 3 | 0 | 0 | 100 | 7 | 3 | 0 | 0 | 100 | 7 | 3 | 0 | 0 | 100 |
| **20** | 3 | 7 | 7 | 3 | 1 | 0 | 100 | 7 | 3 | 1 | 0 | 100 | 7 | 2 | 2 | 0 | 66.6 |
| **Total** | **30** | **70** | **71** | **27** | **3** | **3** | **-** | **74** | **28** | **2** | **5** | **-** | **71** | **27** | **4** | **2** | **-** |
| | | | **Performance = 90%** | | | | | **Performance = 93.3%** | | | | | **Performance = 90%** | | | | |

out of 1500 responses (98.8%, 100% and 99% in three sets of experiments) correctly. The reason for misclassifying 11 responses in Exp. 25-1, 25-3 and 26-1 is that the window used for distance calculations in these response sets was not selected properly. Therefore, some significant parts of several responses were not included in the selected window.

It can be seen from Table 5.4 that Automated MPS found 85% of the SMUPs (17 out of 20) in the first set of experiments and 90% of the SMUPs (18 out of 20) in the second and third sets. Automated MPS could not find one of the two SMUPs correctly in seven experiments. The reason is that in these experiments, the estimation of the probability of observing each combination of the two alternating SMUPs was very different from the number of times it was actually observed. Hence, subtracting the second and third most prevalent groups from the most prevalent group did not give the two correct SMUPs. For example in Exp. 32-3, it can be estimated from Table 5.1 that the most prevalent group is $S1 + baseline$ ($P = 0.8075$). Subtracting the second most prevalent group, baseline ($P = 0.1425$), from the most prevalent group gives $S1$. Similarly, $S2$ is obtained by subtracting the third most prevalent group, $S1+S2+baseline$ ($P = 0.0425$), from the most prevalent group. In the generated response set of this experiment, the most prevalent and the second most prevalent groups are $S1+baseline$ (30 times), and the baseline (18 times). However, the third most prevalent group is $S2+baseline$, (twice), not $S1+S2+baseline$. Therefore, subtracting the third most prevalent group from the most prevalent group gives $S1 - S2$, which does not represent a SMUP.

### 5.5.4 Response Sets with Six Groups

The data was simulated such that two combinations of the selected SMUPs were not generated in order to generate response sets with six groups. Ten data sets with three SMUPs were selected from Table 5.1, and three experiments were performed using each

Table 5.4: Results of applying Automated MPS to incomplete response sets with 3 groups (Two MUs alternate and one group is absent).

| | | | Experiment 1 | | | | | Experiment 2 | | | | | Experiment 3 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Data set** | **Num. of SMUPs** | **Num. of response set groups** | **Num. of detected response set groups** | **Num. of detected SMUPs** | **Num. of misclassified responses** | **Num. of wrong most prevalent groups** | **Performance in finding the SMUPs** | **Num. of detected response set groups** | **Num. of detected SMUPs** | **Num. of misclassified responses** | **Num. of wrong most prevalent groups** | **Performance in finding the SMUPs** | **Num. of detected response set groups** | **Num. of detected SMUPs** | **Num. of misclassified responses** | **Num. of wrong most prevalent groups** | **Performance in finding the SMUPs** |
| **24** | 2 | 3 | 3 | 2 | 0 | 0 | 100 | 3 | 1 | 0 | 0 | 50 | 3 | 2 | 0 | 0 | 100 |
| **25** | 2 | 3 | 3 | 2 | 0 | 1 | 100 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 5 | 100 |
| **26** | 2 | 3 | 4 | 1 | 0 | 5 | 50 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 0 | 100 |
| **27** | 2 | 3 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 0 | 100 |
| **28** | 2 | 3 | 3 | 1 | 0 | 0 | 50 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 0 | 100 |
| **23** | 2 | 3 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 0 | 100 |
| **29** | 2 | 3 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 0 | 100 |
| **30** | 2 | 3 | 3 | 1 | 0 | 0 | 50 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 0 | 100 |
| **31** | 2 | 3 | 3 | 2 | 0 | 0 | 100 | 3 | 1 | 0 | 0 | 50 | 3 | 1 | 0 | 0 | 50 |
| **32** | 2 | 3 | 3 | 2 | 0 | 0 | 100 | 3 | 2 | 0 | 0 | 100 | 3 | 1 | 0 | 0 | 50 |
| **Total** | **20** | **30** | **31** | **17** | **0** | **6** | **-** | **30** | **18** | **0** | **0** | **-** | **30** | **18** | **0** | **5** | **-** |
| | | | **Performance = 85%** | | | | | **Performance = 90%** | | | | | **Performance = 90%** | | | | |

data set. In the first set of experiments, the absent combinations were $S2+S3+baseline$ and $S1+S2+S3+baseline$. The missing combinations were $S1+baseline$ and $S1+S2+S3+baseline$ in the second set and $S1+S2+baseline$ and $S1+S3+baseline$ in the third set.

As explained in Section 4.5.3, when the response set has six groups, Automated MPS can only find two of the three alternating SMUPs. For each experiment, there are two numbers showing the performance. The first number was calculated based on the fact that there are three potential SMUPs in each response set, while the second number represents the performance of Automated MPS assuming that it can find two SMUPs in the best case. Table 5.5 shows the results of applying Automated MPS in 30 experiments. It can be seen that Automated MPS classified 1497 out of 1500 responses (99%) correctly. It found 19 out of 20 SMUPs (95%) in the first set of experiments, and 18 out of 20 SMUPs (90%) in the second and third sets. Considering that there are three potential SMUPs in each experiment, the performance of three sets of experiments is reduced to 63.3% (19 out of 30 SMUPs), 59.9% (18 out of 30 SMUPs) and 59.9% (18 out of 30 SMUPs) respectively. One of the SMUPs is detected incorrectly in Exp. 16-2, 16-3 and 18-3, because an improper window is selected for distance calculations. As described in Section 5.4, the start point of the selected window is the median trimmed average onset of all responses. If the majority of responses have late onsets, the selected window might not include some significant parts of the responses with early onsets and peaks. That is what happens in the above mentioned experiments. However, this is not the case in Exp. 20-2 and 20-1. Automated MPS cannot find one of the SMUPs in these experiments, because the combination of two SMUPs is very similar to the third SMUP, and differentiating between $S3+baseline$ and $S1+S2+baseline$ is difficult.

Table 5.5: Results of applying Automated MPS to incomplete response sets with 6 groups (two groups are absent).

| | | | | **Experiment 1** | | | | | | | **Experiment 2** | | | | | | | **Experiment 3** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Data set** | **Num. of SMUPs** | **Maximum Num. of detectable SMUPs** | **Num. of response set groups** | **Num. of detected response set groups** | **Num. of correctly-detected SMUPs** | **num. of wrong most prevalent groups** | **Num. of wrong detected response groups** | **Num. of misclassified responses** | **Performance in finding the SMUPs (out of 3 SMUPs)** | **Performance in finding the SMUPs (out of 2 SMUPs)** | **Num. of detected response set groups** | **Num. of correctly-detected SMUPs** | **num. of wrong most prevalent groups** | **Num. of wrong detected response groups** | **Num. of misclassified responses** | **Performance in finding the SMUPs (out of 3 SMUPs)** | **Performance in finding the SMUPs (out of 2 SMUPs)** | **Num. of detected response set groups** | **Num. of correctly-detected SMUPs** | **num. of wrong most prevalent groups** | **Num. of wrong detected response groups** | **Num. of misclassified responses** | **Performance in finding the SMUPs (out of 3 SMUPs)** | **Performance in finding the SMUPs (out of 2 SMUPs)** |
| **1** | 3 | 2 | 6 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 6 | 2 | 1 | 0 | 0 | 66.6 | 100 |
| **7** | 3 | 2 | 6 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **8** | 3 | 2 | 6 | 6 | 2 | 2 | 0 | 0 | 66.6 | 100 | 7 | 2 | 0 | 0 | 0 | 66.6 | 100 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **10** | 3 | 2 | 6 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 7 | 2 | 3 | 0 | 1 | 66.6 | 100 | 6 | 2 | 1 | 1 | 0 | 66.6 | 100 |
| **12** | 3 | 2 | 6 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **13** | 3 | 2 | 6 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **15** | 3 | 2 | 6 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 6 | 2 | 0 | 1 | 0 | 66.6 | 100 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **16** | 3 | 2 | 6 | 6 | 2 | 0 | 0 | 0 | 66.6 | 100 | 4 | 1 | 3 | 1 | 0 | 33.3 | 50 | 7 | 1 | 4 | 1 | 1 | 33.3 | 50 |
| **18** | 3 | 2 | 6 | 6 | 2 | 1 | 0 | 0 | 66.6 | 100 | 6 | 2 | 1 | 0 | 0 | 66.6 | 100 | 6 | 1 | 1 | 0 | 0 | 33.3 | 50 |
| **20** | 3 | 2 | 6 | 6 | 1 | 2 | 1 | 0 | 33.3 | 50 | 6 | 1 | 2 | 1 | 0 | 33.3 | 50 | 7 | 2 | 1 | 0 | 1 | 66.6 | 100 |
| **Total** | **30** | **20** | **60** | **60** | **19** | **5** | **1** | **0** | **-** | **-** | **60** | **18** | **9** | **3** | **1** | **-** | **-** | **62** | **18** | **8** | **2** | **2** | **-** | **-** |
| | | | | **Perf. (two SMUPs) = 95%** | | | | | | | **Perf. (two SMUPs) = 90%** | | | | | | | **Perf. (two SMUPs) = 90%** | | | | | | |
| | | | | **Perf. (three SMUPs) = 63.3%** | | | | | | | **Perf. (three SMUPs) = 59.94%** | | | | | | | **Perf. (three SMUPs) = 59.94%** | | | | | | |

### 5.5.5 Response Sets with Five Groups

To generate response sets with five groups, the data was simulated such that three out of eight combinations of the selected SMUPs were not generated in the response set. Ten data sets with three SMUPs were selected from Table 5.1, and three response sets were generated using each data set. In the first set of experiments, the absent combinations were $S2 + S3 + baseline$, $S1 + S3 + baseline$ and $S1 + S2 + S3 + baseline$. The missing combinations were $S3 + baseline$, $S1 + S3 + baseline$ and $S2 + S3 + baseline$ in the second set and $S2 + baseline$ and $S1 + S2 + baseline$ and $S2 + S3 + baseline$ in the third set of experiments.

The results of applying Automated MPS to these 30 generated response sets is shown in Table 5.6. Similar to the cases in which there are six groups in the response set, when the response set has five groups, Automated MPS cannot find all three potential SMUPs (refer to Section 4.5.3). For each experiment, the performance is calculated based on the assumption that there are two SMUPs, as well as the fact that there are three potential SMUPs in the generated response set. It can be seen from Table 5.6 that Automated MPS classified all the responses correctly. 18 out of 20 SMUPs (90%) were detected correctly in the first set of experiments, while 19 out of 20 SMUPs (95%) were detected correctly in the second and third sets of experiments. Considering that there are three potential SMUPs in each experiment, the performance for these experiments is reduced to 59.9% (18 out of 30 SMUPs), 63.3% (19 out of 30 SMUPs) and 63.3% (19 out of 30 SMUPs) respectively. As described in Section 5.4, the end point of the selected window is defined as the median trimmed average peak of the largest 25 responses of the response set plus 10 ms. In Exp. 1-3 and 20-2, some significant parts of several responses are after the selected end point; these parts are not included in the selected window resulting in detecting a wrong SMUP. A bad selection of the start point, which is defined as the median trimmed average onset of all responses, can also lead to detecting wrong SMUPs. That is what happens in Exp. 15-1 and 16-1. It is worth mentioning that selecting a bad window for distance calculations is a bigger problem for simulated experiments compared with real response sets, because in simulated response sets, the SMUPs selected for building the response sets are collected from different sites. Since the responses of different sites have different latencies, the interval between the response peaks or onsets are more than real response sets in which all responses are collected from the same site.

### 5.5.6 Response Sets Collected from Control Subjects

Table 5.7 shows the results of applying Automated MPS to 10 response sets collected from control subjects. In these experiments, Automated MPS found 17 out of 20 SMUPs correctly (85%). Since the number of detected groups in Exp. 5, 7 and 10 are five or six, Automated MPS could not find the third SMUP. The maximum number of detectable SMUPs was 17; all of them were detected correctly by Automated MPS. Table 5.7 shows that Exp. 6, 9 and 10 contained an outlier which was detected and excluded by the operator. Similarly, the response set of Exp. 3 had three outliers, which were detected by the operator. In this experiment and also in Exp. 4 two groups were similar to each other (the difference between them did not represent a SMUP) and were merged together before forming the difference set.

## 5.6 Summary

In this chapter, the results of experiments performed to evaluate the performance of Automated MPS using either simulated response sets or real response sets were reported. The results were discussed and several factors affecting the performance of Automated MPS

Table 5.6: Results of applying Automated MPS to incomplete response sets with 5 groups (Three MUs alternate and three groups are absent).

| | | | | **Experiment 1** | | | | | | | **Experiment 2** | | | | | | | **Experiment 3** | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Data set** | **Num. of SMUPs** | **Maximum Num. of detectable SMUPs** | **Num. of response set groups** | **Num. of detected response set groups** | **Num. of correctly-detected SMUPs** | **num. of wrong most prevalent groups** | **Num. of wrong detected response groups** | **Num. of misclassified responses** | **Performance in finding the SMUPs (out of 3 SMUPs)** | **Performance in finding the SMUPs (out of 2 SMUPs)** | **Num. of detected response set groups** | **Num. of correctly-detected SMUPs** | **num. of wrong most prevalent groups** | **Num. of wrong detected response groups** | **Num. of misclassified responses** | **Performance in finding the SMUPs (out of 3 SMUPs)** | **Performance in finding the SMUPs (out of 2 SMUPs)** | **Num. of detected response set groups** | **Num. of correctly-detected SMUPs** | **num. of wrong most prevalent groups** | **Num. of wrong detected response groups** | **Num. of misclassified responses** | **Performance in finding the SMUPs (out of 3 SMUPs)** | **Performance in finding the SMUPs (out of 2 SMUPs)** |
| **1** | 3 | 2 | 5 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 1 | 4 | 1 | 0 | 33.3 | 50 |
| **7** | 3 | 2 | 5 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **8** | 3 | 2 | 5 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **10** | 3 | 2 | 5 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 1 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **12** | 3 | 2 | 5 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **13** | 3 | 2 | 5 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **15** | 3 | 2 | 5 | 5 | 1 | 2 | 1 | 0 | 33.3 | 50 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **16** | 3 | 2 | 5 | 5 | 1 | 1 | 0 | 0 | 33.3 | 50 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **18** | 3 | 2 | 5 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **20** | 3 | 2 | 5 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 | 5 | 1 | 1 | 0 | 0 | 33.3 | 50 | 5 | 2 | 0 | 0 | 0 | 66.6 | 100 |
| **Total** | **30** | **20** | **50** | **50** | **18** | **3** | **1** | **0** | **-** | **-** | **50** | **19** | **2** | **0** | **0** | **-** | **-** | **50** | **19** | **4** | **1** | **0** | **-** | **-** |
| | | | | **Perf. (two SMUPs) = 90%** | | | | | | | **Perf. (two SMUPs) = 95%** | | | | | | | **Perf. (two SMUPs) = 95%** | | | | | | |
| | | | | **Perf. (three SMUPs) = 59.94%** | | | | | | | **Perf. (three SMUPs) = 63.3%** | | | | | | | **Perf. (three SMUPs) = 63.3%** | | | | | | |

Table 5.7: Results of applying Automated MPS to real response sets.

| **Data set** | **Num. of SMUPs** | **Num. of detected SMUPs** | **Num. of correctly-detected SMUPs** | **Num. of response set groups** | **Num. of detected response set groups** | **Num. of detected response set groups after merging** | **Num. of most prevalent groups** | **Num. of correctly-detected most prevalent groups** | **Num. of misclassified responses** | **Performance in finding the SMUPs** |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 1 | 1 | 1 | 2 | 2 | 2 | - | - | - | 100 |
| **2** | 2 | 2 | 2 | 4 | 4 | 4 | 4 | 4 | 0 | 100 |
| **3** | 2 | 2 | 2 | 4 | 8 | 4 | 4 | 4 | 7 | 100 |
| **4** | 1 | 1 | 1 | 2 | 3 | 2 | - | - | - | 100 |
| **5** | 3 | 2 | 2 | 5 | 5 | 5 | 4 | 2 | 0 | 100 |
| **6** | 2 | 2 | 2 | 4 | 5 | 5 | 4 | 2 | 1 | 100 |
| **7** | 3 | 2 | 2 | 5 | 5 | 5 | 4 | 3 | 8 | 66.6 |
| **8** | 2 | 2 | 2 | 3 | 3 | 3 | - | - | - | 100 |
| **9** | 1 | 1 | 1 | 3 | 3 | 3 | - | - | - | 100 |
| **10** | 3 | 2 | 2 | 5 | 6 | 5 | 4 | 4 | 0 | 33.3 |
| **Total** | **20** | **17** | **17** | **37** | **44** | **43** | **24** | **19** | **16** | **-** |

were explained briefly. The next chapter will describe these factors in more detail.

# Chapter 6

# Discussion

In Chapter 5, the ability of Automated MPS to detect the constituent SMUPs of a response set was reported by presenting the results of several experiments completed to evaluate its performance. The first part of this chapter explains why Automated MPS might not perform well in some special cases. The factors with a negative impact on the performance of Automated MPS were reviewed very shortly in Chapter 5; here, they are explained in more detail. Then, the advantages and disadvantages of Automated MPS over other MUNE techniques are explained. Finally, Automated MPS is compared with two other similar MUNE techniques, namely MPS and MUESA.

## 6.1 Factors Affecting the Performance of Automated MPS

Assuming that the electrode does not move relative to the muscle during response set collection, and only one to three SMUPs alternate, Automated MPS should be able to detect all existing SMUPs. However, if one of the following special scenarios happens, one or more detected SMUPs might be incorrect.

-**The noise level is high:** If the noise level of the waveforms of a response set is high compared with SMUPs sizes, differentiating between certain clusters becomes difficult. Since the response set clusters are the input of the second step of Automated MPS, if the responses are not grouped correctly, the extracted SMUPs (some or all of them) will also be wrong.

- **The selected window does not include some significant parts of some responses:** As explained in Section 5.4, selecting a proper portion of the recorded waveforms for distance calculations is necessary for finding correct single SMUPs; this portion should be limited to the significant parts of the responses. If the start point of the selected window is very early (the onsets are determined incorrectly), or the end point is too late (the determined end points are incorrect), and the selected window is very large, some non-important parts of the responses that might contain significant ups and downs, such as stimulus artifact or an F-response, might be included in distance calculations resulting in some errors in clustering the waveforms of the response set or the difference set. On the other hand, selecting a very small window also has a negative impact on the performance of Automated MPS. If the start point of the selected window is too late (the determined onsets are correct, but the onsets of the majority of the responses are too late), some significant parts of the responses with early onsets are not included in the selected window. The same thing occurs if the end point of the selected window is too early (the negative peaks of half of the largest responses are too early). The second problem is more common in simulated response sets, because the constituent SMUPs of simulated response sets are collected from different sites with different latencies, while in real response sets all responses are collected from the same site; therefore, the peaks and onsets of all responses are usually close together (assuming that the landmarks are detected correctly).

- **An outlier is detected as a singleton response group and vice versa:** An observed singleton in a response set can be the result of voluntary contraction or movement of the electrodes. On the other hand, a singleton can be one of the combinations of the $n$ SMUPs that has been observed once, because it has a small probability of observation. In the current version of Automated MPS, the operator should visually decide whether a singleton is an outlier or not. This decision making depends on the experience of the operator and can affect the results.

- **The combination of two SMUPs is very similar to the third SMUP:** In experiments performed in Chapter 5, there are several cases in which the Automated MPS finds a wrong SMUP, because two response groups, the combination of two SMUPs ($S1 + S2 + baseline$) and the third SMUP ($S3 + baseline$), are very similar to each other. The same problem occurs when one of the SMUPs is so small compared with the second SMUP. In this case the combination of these two SMUPs is very similar to the second SMUP (the larger SMUP).

The above mentioned scenarios can affect the results of Automated MPS either when a full response set is observed, or some of the combinations are absent. The following scenarios exclusively happen when the response set is incomplete.

- **Response sets with three groups:** When the number of alternating MUs is two, and the number of observed response groups is three, Automated MPS can find one or both SMUPs by subtracting the second and third most prevalent groups of the response set from the most prevalent group (refer to Section 4.5.2). If the observed most, second most and third most prevalent groups of the response set are not the same as the expected most, second most and third most prevent groups, one or both detected SMUPs might be wrong.

- **Response sets with three or seven groups:** If one of the combinations of three alternating SMUPs is absent in the response set, it is not possible to find the three SMUPs by comparing the $2n$ most prevalent groups of the difference set with the response set groups, because it is possible that the absent group is one of the SMUPs. According to Section 4.5.1, if one of the responses is absent, Automated MPS looks for the three most prevalent groups of the difference set that have their negative peak before their positive peaks. To use this strategy, it should be assumed that the negative peak always comes before the positive peak. Although this assumption is true most of the time, it is possible that the positive peak of an SMUP comes before its negative peak. If this happens, using the above decomposition technique leads to detecting the inverse of the actual SMUP as an SMUP. The same problem occurs when the number of response set groups is three. Automated MPS subtracts the second and third most prevalent response from the most prevalent response to find the two SMUPs. If the positive peaks of the resulting waveforms come before their negative peaks, Automated MPS inverts them.

- **Response sets with five or six groups:** If two or three out of eight possible combinations of the responses of three alternating MUs are not observed in the response set, Automated MPS can usually find two SMUPs. However, there are some special cases in which Automated MPS is not able to detect one or both SMUPs correctly. These special scenarios are summarized in the following paragraphs for cases in which the number of response set groups is five. The same problems can occur if there are six groups in the response set.

- When the response set group has only five response groups (three MUs alternate, and three combinations are missed), Automated MPS can find two SMUPs out of three if four out of five observed groups represent four different combinations of two of the three alternating SMUPs. However, there are special cases in which no SMUP can be obtained from a response set with five groups. For example, if the observed response set includes the baseline, $S1+S2+baseline$, $S1+S3+baseline$, $S2+S3+baseline$, and $S1+S2+S3+baseline$, Automated MPS cannot extract any correct SMUP because there is not four groups that can represent the four combinations of any two SMUPs.

- If certain combinations are not observed in the response set, Automated MPS can find four groups that apparently represent the four combinations of two SMUPs. However, one of the detected SMUPs is actually the combination of two SMUPs. For example, assume that the response set includes baseline, $S1+baseline$, $S2+S3+baseline$, $S1+S2+S3+baseline$, and $S2+baseline$, Automated MPS chooses the first four groups as the best four groups for decomposition. Decomposing these four groups results in detecting $S1$ and $S2+S3$ as two SMUPs. Since $S2+S3$ is not the response of a single SMUP, including it in the calculation of the mean-SMUP leads to the overestimation of the mean-SMUP and underestimation of the number of MUs. This case happens when the firing thresholds of $S2$ and $S3$ are very close to each other; therefore, there is no observed response in which one of the SMUPs activates, but the other one does not activate. The same problem and the same mistake can also happen in MPS.

- If one of the missing responses is the baseline response, it is possible that Automated MPS detects another observed group (the one with the lowest energy) as the baseline response. For example, if the response set includes $S1+baseline$, $S2+baseline$, $S3+baseline$, $S2+S3+baseline$ and $S1+S2+S3+baseline$, Automated MPS chooses the first four groups for decomposition, and detects $S1+baseline$ as the baseline response. When the four most prevalent groups of the difference set are compared to the selected four groups, ($S2+baseline$)-($S1+baseline$) $= S2-S1$ and

$(S3 + baseline)$-$(S1 + baseline)$=$S3 - S1$ are detected as two out of three existing SMUPs. In this case, both detected SMUPs are incorrect.

## 6.2 Comparing Automated MPS with Other MUNE Techniques

As explained in Chapter 2, Automated MPS was developed to overcome the deficiencies of the MPS [17] and MUESA [34] techniques. In this section, the pros and cons of Automated MPS are discussed. Then, Automated MPS is compared with MPS and MUESA to explain how the Automated MPS method improves upon these techniques.

### 6.2.1 Automated MPS Advantages

Advantages of Automated MPS as compared with other existing MUNE methods include:

- It can handle alternation. Alternation is the main problem of many MUNE techniques. These techniques work based on the assumption that each observed signal is the result of the activation of a new single MU. Automated MPS however, can be performed when a single MU is activated, as well as when two or three MUs are probabilistically active.

- Since in Automated MPS constant-intensity stimulations are applied to the nerve, a limited number of MUs are probabilistically active after each pulse, and possible changes in electrophysiological response can be examined. Most MUNE techniques assume that there is no change in the response of each MU during the whole procedure. However, in some diseases such as ALS, the amplitude of the CMAP changes if the stimulation is repeated several times.

- Unlike some MUNE techniques that estimate a mean-SMUP, Automated MPS calculates a mean-SMUP by averaging several real SMUPs.

### 6.2.2 Automated MPS Disadvantages

Automated MPS has the following deficiencies compared with other MUNE techniques.

- Although the operator working with Automated MPS does not need as much skill as an MPS operator, he/she should be experienced enough to detect how many MUs are alternating at each time. He/she should also be able to set the stimulation such that one, two or three MUs alternate.

- If certain combinations of $n$ SMUPs are missing (refer to Section 6.1), Automated MPS cannot obtain any accurate SMUPs from the observed response set and a new set should be collected. As the number of sets required for calculating the mean-SMUP increases, the patient comfort decreases.

- It cannot be applied to proximal muscles.

### 6.2.3 Comparing Automated MPS and MUESA

Automated MPS is similar to MUESA in some steps. In both methods, several responses are collected from each site in which one, two or three MUs alternate. The responses are sorted and clustered; then the response set is decomposed to extract the constituent SMUPs. The major difference between Automated MPS and MUESA is the decomposition technique they employ. MUESA decomposes the response set based on the size of the responses if a full set is observed, or based on how many times each combination is observed either when a full response set is observed or one combination is absent. Automated MPS finds the $n$ SMUPs by comparing the $2n$ most prevalent groups of the difference set with response set groups if the response set is full or one of the combinations is absent; other techniques are used if more than one combination is absent. The main advantage of MUESA compared with Automated MPS is its simplicity. However, Automated MPS has the following advantages over MUESA.

- MUESA cannot deal with phase cancellation. The decomposition technique used in MUESA is based on the assumption that the size of the combination of two SMUPs is always larger than the size of each of them. If there is some phase cancellation as the waveforms summate, this assumption is not true and the combination of two SMUPs might be smaller than each SMUP. Automated MPS however can handle phase cancellation in situations where the number of response set groups is four, five, six or eight.

- MUESA employs another decomposition approach based on the estimation of firing rates of the alternating MUs. This approach only works correctly when the expected order of prevalence of the responses is the same as the order of their prevalence in the observed response set. It can be seen from the results of Automated MPS in Table 5.4 of Section 5.5.3 that the number of times each response is observed in a response set can be very different from the expected numbers. If this happens, it is possible that the expected most prevalent response is the second or third most prevalent response in the collected response set. This leads to extracting an incorrect SMUP. The decomposition technique in Automated MPS does not suffer from this problem, except in cases where the number of response set groups is three. When there are three groups in the response set, Automated MPS uses the same approach as MUESA and therefore, suffers from the same problem).

- MUESA does not propose a decomposition technique when the number of response set groups is five or six (i.e., more than one combination is missed), as opposed to Automated MPS in which usually one or two SMUPs can be detected when two or three combinations are missed.

It is worth mentioning that although a thesis and paper were published about MUESA, some important details of this method were not available. Therefore, implementing MUESA and a direct comparison between MUESA and Automated MPS was not possible in this work.

### 6.2.4 Comparing Automated MPS and MPS

MPS [17] is an alternation-free MUNE technique in which the motor nerve is stimulated at several sites, and one SMUP (the response of the first activated MU) is collected at each site. The main disadvantage of MPS is that the operator needs to have considerable skills and experience to collect a sufficient number of SMUPs in a reasonable time (about 20 min), and recognize alternation or other errors that prevent the identification of the stimulation of a single MU. The operator working with Automated MPS needs less skill and experience compared with an MPS operator. For example, the operator does not need to be careful to avoid alternation, because Automated MPS can handle alternation. Moreover, in Automated MPS one, two or sometimes three SMUPs are collected from each stimulus site. Therefore, enough SMUPs (about ten) can be collected from three or four sites, as opposed to the MPS method in which, about 10 sites need to be stimulated.

# Chapter 7

# Conclusions and Future Work

## 7.1 Conclusions

In this thesis, a new decomposition technique for finding single SMUPs was introduced. In this technique, called Automated MPS, a muscle is stimulated with a train of constant-intensity pulses, and 50 surface detected waveforms are recorded while one, two or three MUs alternate. Each collected waveform is one of the $2^n$ possible combinations of $n$ alternating MUs. The waveforms of the response set are clustered by forming a minimum spanning tree and cutting it with a flexible threshold. Then, if there are similar groups in the response set (i.e., if the difference between two groups cannot represent a SMUP physiologically), they are merged together. Moreover, if there is any singleton in the response groups, the operator decides whether it is an outlier caused by patient or electrode movement, or if it is one of the combinations, which has been observed only once. Then, depending on the number of response set groups, a decomposition procedure is applied to the response set to obtain the $n$ constituent SMUPs (No SMUP can be obtained from a response set with one group). A new decomposition technique for cases with four, seven and eight groups is introduced in this thesis. It can also be applied to response sets with five or six groups provided that one or two groups are excluded. In this decomposition method, each response group is subtracted from the other response groups forming a set, called the difference set. The waveforms of the difference set are clustered; then, the $2n$ most prevalent groups of the difference set are compared to the response set groups to obtain the $n$ SMUPs. On the other hand, if there are two groups in the response set, finding the only SMUP is straightforward, and if there are three groups, the MUESA decomposition technique, which is based on the estimation of the firing rates of the SMUPs, is used.

The experiments performed in Chapter 5 used both real response sets and simulated response sets to evaluate the performance of Automated MPS. Most of the experiments used simulated response sets; however, these simulated response sets resembled real data. Different combinations of $n$ SMUPs were generated using $n$ real SMUPs. A variable noise was also added to the simulated waveforms to make it as real as possible. The generated response sets were first clustered. Then, these clusters were used as the input to Automated MPS for decomposition. The results of our experiments were satisfactory showing that the proposed decomposition technique works well. The performance of Automated MPS was much higher when the response sets were complete. Several factors that could disturb the results of clustering the response set and difference set were described in Section 6.1, as well as some other factors which were not observed in the experiments, but are possible in real data sets.

In Chapter 2, we claimed that Automated MPS is an improvement over MPS and MUESA. These three techniques were compared with each other in Chapter 6. Automated MPS has two main advantages over MPS. First, working with Automated MPS is much easier for the operator, because the operator needs less experience. Our experiments prove this fact. Second, several SMUPs can be obtained from one site; therefore, fewer sites are needed to obtain enough SMUPs for calculating a mean-SMUP. Furthermore, Automated MPS improves the decomposition techniques used in MUESA; because it can detect more accurate SMUPs in more difficult situations.

In conclusion, it is confidently reported that Automated MPS can obtain enough SMUPs much easier than MPS and much more accurately than MUESA. However, further improvements can still be made to the clustering part, as well as the decomposition part of Automated MPS.

## 7.2 Future Works

The following recommendations might be useful to build the road for possible future studies.

- The focus of this thesis was to use pattern recognition techniques to find a decomposition tool for obtaining single SMUPs. Therefore, in our experiments, we did not collect a maximal-CAMP, nor estimate the number of MUs in the muscle. Further experiments should be performed with real data; the mean-SMUP should be calculated, a maximal-CMAP should be collected, and the number of MUs should be estimated to prove that Automated MPS can give a good estimation of the number of MUs of a muscle. The best way to do this comparison is to estimate the number of MUs in muscles of the same patients using Automated MPS and another reliable MUNE technique such as MPS, and compare the results.

- As discussed in Chapter 6, if the baseline response (i.e., the response when none of $n$ alternating MUs are activated) is not observed in the response set, Automated MPS might have some difficulties in finding the SMUPs. Further studies should be done to find a way for estimating the baseline response even if it is not observed in the response set.

- In the current version of Automated MPS, the operator should visually decide about the singleton response set groups. The decision made by an operator depends on the amount of experience he/she has. Further research should be done to enable Automated MPS to make this decision automatically.

- To the best of author's knowledge, up to now, there is no decomposition technique for finding the third SMUP when the number of response set groups is five or six. Further studies should be done to find a method for detecting the third SMUP.

- The ability of reporting an error in cases in which the decomposed SMUPs cannot reproduce the observed response set should be added to Automated MPS.

# Bibliography

[1] Kimball's biology pages: Muscles, Aug. 2007. URL = http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/M/Muscles.html.

[2] Motor unit and recruitment, Aug. 2007. URL = www.holycross.edu/departments/biology.

[3] J.P. Ballantyne and S. Hansen. A new method for the estimation of the number of motor units in a muscle: Control subjects and patients with myasthenia gravis. *J Neurol Neurosurg Psychiatry*, 37:907–915, 1974.

[4] J.V. Basmajian and C.J. De Luca. *Muscles Alive; Their functions revealed by electromyography.* Williams and Wilkins, USA, fifth edition, 1985.

[5] J. Bergmans. Physiologic observation of single human nerve fibres. *Electromyography and Clinical Neurophysiology*, 2:89–127, 1973.

[6] J. Bethlem. *Myopathies.* North-Holland Biomedical Press, 1980.

[7] M.B. Bromberg. Motor unit estimation: reproducibility of the spike-triggered averaging technique in normal and als subjects. *Muscle and Nerve*, 16(5):466–471, 1993.

[8] M.B. Bromberg, editor. *Motor Unit Number Estimation, Supplement to Clinical Neu-*

*rophysiology Series*, volume 55. ELSEVIER, Snowbird, Utah, USA, 2001. Proceedings of the First International Symposium on MUNE.

[9] W.F. Brown. A method for estimating the number of motor units in thenar muscles and the changes in motor unit count with ageing. *J Neurol Neurosurg Psychiatry*, 35:845–852, 1972.

[10] W.F. Brown, M.J. Strong, and R. Snow. Methods for estimating numbers of motor units in biceps-brachialis muscles and losses of motor units with aging. *Muscle and Nerve*, 11(5):423–432, 1988.

[11] J.R. Daube. Estimating the number of motor units in a muscle. *Clinical Neurophysiology*, 12:585–594, 1995.

[12] J.R. Daube. Mune by statistical analysis. *Clinical Neurophysiology*, 55:51–71, 2003.

[13] T. Doherty, K. Chan, , and W. Brown. *Neuromuscular function and disease: Peripheral Motor System.* W.B. Saunders comp., 2002.

[14] T. Doherty, D. Stashuk, , and W.F. Brown. Axon excitability in motor unit number estimation. *Clinical Neurophysiology*, 55:17–21, 2003.

[15] T. Doherty, D. Stashuk, and W.F. Brown. Mune: Multiple point stimulation and f-response methods. *Clinical Neurophysiology*, 55:31–40, 2003.

[16] T.J. Doherty and W.F. Brown. The estimated numbers and relative sizes of thenar motor units as selected by multiple point stimulation in young and older adults. *Muscle and Nerve*, 16:355–366, 1993.

[17] T.J. Doherty and W.F. Brown. *Clinical Neurophysiology and Neuromuscular Disease.*, pages 274–290. W.B. Saunders, Philadelphia, USA, 2002. Motor Unit Number Estimation: Methods and Applications.

[18] M.S. Emrani. Relationships between motor unit anatomical characteristics and motor unit potential statistics in healthy muscles. Master's thesis, Dept. of Systems Design Engineering, University of Waterloo, 2005.

[19] V. Galea, H. Bruin, R. Cavasin, , and A.J. McComas. The numbers and relative sizes of motor units estimated by computer. *Muscle and Nerve*, 14:1123–1130, 1991.

[20] A.C. Guyton and J.E. Hall. *Text Book of Medical Physiology.* W.B. Saunders comp., 1996.

[21] A.K. Jain and R.C. Dubes. *Algorithms for Clustering Data.* Prentice Hall, 1988.

[22] A.K. Jain, M.N. Murty, and P.J. Flynn. Data clustering: a review. *ACM Computing Surveys*, 31(3):264–323, 1999.

[23] H.A. Kadrie, S.K. Yates, H.S. Milner-Brown, and W.F. Brown. Multiple point electrical stimulation of ulnar and median nerves. *J Neurol Neurosurg Psychiatry*, 39:973–985, 1976.

[24] G. Karypis, E.H. Han, and V. Kumar. Chameleon hierarchical clustering using dynamic modeling. *IEEE Computer*, 32(8):68–75, 1999.

[25] S.Y. Lu and K.S. Fu. A sentence-to sentence clustering procedure for pattern analysis. *IEEE Trans. Syst. Man Cybern.*, 8:381–389, 1978.

[26] R. Luescher, P. Ruenzel, and E. Henneman. How the size of motoneurons determines their susceptibility to discharge. *Nature*, 282:859–861, 1979.

[27] J.B. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1967.

[28] A.J. McComas, P.R.W. Fawcett, and M.J. Campbell. Electrophysiological estimation of the number of motor units within a human muscle. *Journal of neurology and psychiatry*, 34(2):121–131, 1971.

[29] F. Jahanmiri Nezhad. Motor unit firing pattern analysis for adaptive emg signal decomposition. Master's thesis, Dept. of Systems Design Engineering, University of Waterloo, 2006.

[30] M.G. Omran. *Particle swarm optimization methods for pattern recognition and image processing.* PhD thesis, Univ. of Pretoria, 2005.

[31] C.P. Panayiotopoulos, S. Scarpalezos, and T. Papapetropoulos. Electrophysiological estimation of motor units in duchenne muscular dystrophy. *J Neurol Sci.*, 23:89–98, 1974.

[32] R.C. Prim. Shortest connection networks and some generalizations. *Bell System Technical Journal*, 36:1389–1401, 1957.

[33] P. G. Ridall, A. Pettitt, R. Henderson, and P. McCombe. Motor unit number estimation-a bayesian approach. *Biometrics*, 62(4):1235–1250, 2006.

[34] M. Slawnych, C. Laszlo, and C. Herschler. Motor unit estimates obtained using the new muesa method. *Muscle Nerve.*, 19(5):626–636, 1996.

[35] D. Stashuk. Decomposition and quantitative analysis of clinical electromyographic signals. *Medical Engineering and Physics*, 21:389–404, 1999.

[36] D.W. Stashuk and W.F. Brown. *Neuromuscular function and disease: Basic, clinical and diagnostic aspects.*, volume 1. Elsevier Science, Pennsylvania, USA, 2002.

[37] D.W. Stashuk, T. Doherty, A. Kassam, and W.F. Brown. Motor unit number estimates based on the automated analysis of f- responses. *Muscle and Nerve*, 17:881–890, 1994.

[38] D.W. Stashuk, T.J. Doherty, and W.F. Brown. Mune using decomposition-enhanced spike-triggered averaging. *Clinical Neurophysiology*, 55:108–121, 2003.

[39] M.J. Strong, W.F. Brown, A.J. Hudson, and R. Snow. Motor unit estimates in the biceps-brachialis in amyotrophic lateral sclerosis. *Muscle and Nerve*, 11(5):415–422, 1988.

[40] F. Wang and P.J. Delwaide. Number and relative size of thenar motor units estimated by an adapted multiple point stimulation method. *Muscle and Nerve*, 18(9):969–979, 1995.

[41] F.C. Wang, O. Bouquiaux, and V. Pasqua. Adapted multiple point stimulation mune technique. *Clinical Neurophysiology*, 55:41–45, 2003.

[42] F.C. Wang and P.J. Delwaide. An adapted multiple point stimulation (amps) method for estimating the number and size of motor units. *Electroencephalography and Clinical Neurophysiology, Electromyography and Motor Control*, 97(4):164–164(1), 2003.

[43] H. Zhang. A note on fuzzy clustering. URL = http://www.cse.uconn.edu/samples/Zhang report.pdf.